

NON-SLAVIC LANGUAGES OF THE USSR

Papers From The Fourth Conference

Edited by

Howard I. Aronson

Slavica Publishers, Inc.

NON-SLAVIC LANGUAGES OF THE USSR

Papers From The Fourth Conference

NON-SLAVIC LANGUAGES OF THE USSR

Papers From The Fourth Conference

Edited by

Howard I. Aronson

Slavica Publishers, Inc.

Slavica publishes a wide variety of scholarly books and textbooks on the languages, peoples, literatures, cultures, folklore, history, etc. of the former USSR and Eastern Europe. For a complete catalog of books and journals from Slavica, with prices and ordering information, write to:

Slavica Publishers, Inc.
PO Box 14388
Columbus, Ohio 43214

ISBN: 0-89357-250-0.

Each article copyright © 1994 by its author. All rights reserved.

All statements of fact or opinion are those of the authors and do not necessarily agree with those of the publisher, which takes no responsibility for them.

Printed in the United States of America.

Contents

Foreword.	5
Introduction.	9
Howard I. Aronson	
Paradigmatic and syntagmatic subject in Georgian.	13
Roland Bielmeier	
On Iranian influence in old Georgian.	34
J. C. Catford	
Vowel Systems of Caucasian languages.	44
John Colarusso	
How to describe the sounds of the Northwest Caucasian languages.	61
Victor Friedman	
Assertive verb forms in Lak.	114
Zbigniew Gołąb	
Prehistoric contacts between Ossetic and Slavic.	120
Alice C. Harris	
On the history of relative clauses in Georgian.	130
Dee Ann Holisky	
Notes on Auxilliary verbs in Tsova-Tush (Batsbi).	143
Johanna Nichols	
The structure of the Nakh-Daghestanian verb root and verb stem.	160
Alfred G. Paludis	
The subjunctive in classical Armenian: significant differences between Eznik and Elišē.	180
K. H. Schmidt	
Class inflection and related categories in the Caucasus.	185
Wolfgang Schulze-Fürhoff	
Tracing aspect coding techniques in the Lezgian languages.	193

David Testen

The correspondence: Scythian Βαστακας = Ossetian bästä. 209

Kevin Tuite

Syntactic subject in Georgian. 218

Robert Austerlitz

Gilyak internal reconstruction, 3: ligneous matter. 229

Donald L. Dyer

Moldavian linguistic realities. 234

Rachel Lehr

Complex infinitives and other deverbal nominals in Tajik. 254

Jules Levin

Stressing freely in Lithuanian and Russian. 264

Roy Andrew Miller

The original geographic distribution of the Tungus languages. . . . 272

Stefan Pugh

Observations on the Russian component in Karelian. 298

Steven Young

The scope of Saussure's law in colloquial Lithuanian. 304

Foreword

The Fourth International Conference on the Non-Slavic Languages of the USSR was held at the University of Chicago in spring, 1985. Many of the papers in this volume were presented in preliminary form at that conference.

The papers dealing with the languages of Caucasia presented at the Conference were dedicated to Aḡaḡi Šanije on the occasion of his 98th birthday. We had hoped to present him with the published version of these papers and other papers that could not be presented at the Conference in 1987, on the occasion of his one hundredth birthday. But, on 29 March 1987, just one month after the celebration of his centenary on 26 February, Aḡaḡi Šanije passed away, and so this collection has become, instead, a memorial volume, dedicated to the lasting memory of a great scholar.

I must express my gratitude to Bill J. Darden, co-director of our conference, to David Testen, and to the participants in the conference. Special thanks go to Charles Gribble, Nancy T. Buterbaugh, and Slavica Publishers for preparing the copy of this volume, an especially difficult and complex task.

I should like to personally dedicate this volume to the memory of Aḡaḡi Šanije and to the memory of my colleague and friend at the University of Chicago, Zbigniew Goḡab.

Howard I. Aronson



სტუდიები კავკასიის
ენათმეცნიერებაში

STUDIES IN THE
LINGUISTICS OF CAUCASIA

IN MEMORY OF

AKAŖI ŠANIJE

1887-1987



ჯანსუღი გიორგი

In Memoriam
AĶAĶI ŖANIJE

26.II.1887–29.III.1987

Few scholars and fewer linguists have played as major a role in the history of their countries as did AĶaĶi Ŗanije (Shanidze). A founder of Tbilisi University, a founder of the Georgian Academy of Sciences, first to defend a doctoral dissertation at Tbilisi University, a codifier of the modern Georgian literary language, folklorist, philologist, linguist, *BaĶoni AĶaĶi* deservedly became one of the great heroes of the Georgian people, a nation not lacking in heroes. But Ŗanije's significance extends far beyond the borders of his native land. For many of us outside of Georgia, his brilliant elucidation of the seeming chaos of Georgian grammar opened the door to the study of Georgian and the other Kartvelian languages, thereby making the wonders of Georgian history, literature, and all aspects of its culture accessible to us. For others of us, knowledge of the Georgian language, in its turn, helped to unveil the mysteries of the other languages and cultures of that "mountain of languages," the Caucasus. Ŗanije's influence on Kartvelian studies and general linguistic studies continues to be manifested in the works of these scholars both inside and outside of Georgia who felt his powerful influence, either directly or indirectly.

There is scarcely an area of Georgian and Kartvelian language study that has not benefited from Ŗanije's scholarship. Not just a leading philologist, he was also a true linguist in the most modern sense, a student of, among others, Baudouin de Courtenay and ŖĶerba. Structuralist linguists reading Ŗanije's works immediately find themselves in the presence of a kindred spirit.

Ŗanije was an exceptionally productive scholar. A bibliography¹ of his publications in the seventy year period from 1906 through 1976 contains over 760 entries – and it must be remembered that Ŗanije's scholarly productivity continued unabated until his death more than ten years later. The breadth of his interests is astonishing: folklore, history, literature,

bibliography, pedagogy, Rustavelology, and in the domain of linguistics: dialectology, lexicology and lexicography, etymology, paleography, language codification. But it is not just the quantity of his research. The quality of Šanije's contributions is striking. One need just compare the grammars of Georgian written before Šanije with his to see what clarity he brought to the description of the structure of the language.

Šanije's major contributions were to the history and structure of Georgian and the Kartvelian languages. Unquestionably our current understanding of the history and structure of Georgian results from his scholarly investigations in these areas. Among the many terms and concepts that he introduced to the study of Georgian are such essential notions as *xanmeṭi* and *haemeṭi* texts, *version*, *screeve*. It is Šanije who first proposed that Mingrelian and Laz be viewed as dialects of one language. The influence of his *Kartuli gramaṭika, I. Morpologia* [Georgian grammar, I. Morphology; Tbilisi, 1930] and the monumental *Kartuli gramaṭikis sapujvlebi* [Fundamentals of Georgian grammar; Tbilisi, 1953] can be seen in every subsequent analysis of Georgian, both inside and outside of Georgia. It is apparent in the grammatical studies of Rudenko, Vogt, and Tschenkéli, which in turn helped to make the structure of Georgian accessible to those who did not know Georgian. And his influence is still apparent in the works of the younger generations of Kartvelologists both inside and outside of Georgia.

For all practical purposes, Šanije is the codifier of the contemporary Georgian literary language. His Georgian language textbooks have introduced generations of Georgian pupils to the normative grammar of their native language.

The study of the history of Georgian language, the grammar of Old Georgian, and the study of Old Georgian literary texts, owe an immeasurable debt to Šanije's contributions to these fields. His *Jveli kartuli enis gramaṭika* [Grammar of Old Georgian;² Tbilisi, 1976], volume 18 in the outstanding series *Jveli kartuli enis ḡatredris ṡromebi* [Works of the Chair of Old Georgian Language], founded by Šanije, is a succinct but thorough survey of the grammar of the oldest Georgian texts. He is the editor of numerous definitive editions of Old Georgian texts. Simply stated, there is no area of Georgian linguistics and philology where Aḡaḡi Šanije has not left his mark.

One should not omit mention of Šanije's contributions to the study of Caucasian Albanian. To him belongs the first description and analysis of the alphabet of this Caucasian Christian nation.

With these papers, we, non-Georgian scholars working in the fields of Caucasian linguistics, express our admiration, gratitude, and respect to a great scholar who has had an influence on all of us. We hope that these papers will serve as a fitting tribute to the memory of Aḳaḳi Šanije and testify to his enduring influence on our fields of study.

Howard I. Aronson
Chicago, June 1994

Notes

¹Jijiguri [Dzidziguri], Šota, chief editor. 1977. *Aḳaḳi Šanije: biobibliograpia/Akakij Šanidze: biobibliografija*. Tbilisi: Mecniereba. In addition to listing his publications, the volume contains brief biographical sketches of Šanije in Georgian (pp. 5–24), Russian (pp. 25–44), and German (pp. 45–55).

²Happily, an excellent translation by Heinz Fähnrich of this grammar exists: Schanidse, Akaki, 1982. *Altgeorgisches Elementarbuch, 1. Teil. Grammatik der altgeorgischen Sprache*. Tbilissi: Staatsuniversität Tbilissi, Schriften des Lehrstuhls für altgeorgische Sprache, 24.

Paradigmatic and Syntagmatic Subject in Georgian

Howard I. Aronson

Paradigmatic and syntagmatic, case and subject. It is a well known fact of structuralist linguistics that a given *case* form, such as the Georgian ergative, e.g., *monadire-m* 'hunter-ERG' exists as a consequence of its being opposed to the remaining case forms, members of a given paradigmatic set; e.g.:

NOM	monadire-
ERG	monadire-m
DAT	monadire-s
GEN	monadir -is
INSTR	monadir -it
ADV	monadire-d

As Stankiewicz has pointed out (1962:6-7), the members of a paradigm such as that above are in a relationship of mutual implication; the presence of the ERG implies the presence of all the remaining cases, as their presence implies the presence of an ERG.¹ The very definition of the ERG in Georgian must be phrased in terms of its differentness from the remaining cases of the system.

The notion of *subject*, on the other hand, although possibly existing in a paradigmatic relationship with other notions such as direct object and indirect object, etc., is primarily a syntagmatic category. It marks a relationship between a certain specified noun phrase and the verb in a language having such categories, i.e., is a syntactic category.

The notion of subject with which we shall be operating here is a purely grammatical one, one that is language specific and one that must be defined in terms of the structure of the language in which it occurs. There is no need to assume, in fact, that subject is a universal grammatical category.² The noun phrase selected as the grammatical subject can generally be regarded as an invariant of the sentence,³ the unmarked member of a hierarchy including such other syntagmatic units as the direct object and the indirect object. Establishing a grammatical category of subject generally greatly simplifies the stating of the rules of

syntax; its use therefore should contribute to overall economy of the description.

I must emphasize that I am here attempting to give a chapter in a grammar that attempts to account for how the addressee is able to decode the message transmitted by the speaker-encoder. This is, then, not an approach that attempts to account for how the message is encoded, i.e., for the behavior of the speaker. Therefore we shall look for information on how to determine such grammatical information as transitivity and subject/object relations solely on the surface level.

In a goodly number of languages, a grammatical subject can only occur in one given case, e.g., the NOM in Russian. In such languages all grammatical subjects are NOM but not all NOMs are grammatical subjects, the notion of grammatical subject being syntagmatic. So, in Russian utterances the NOM can function as a vocative, as a predicate nominative, or simply to name ("Nennfunktion") in addition to functioning as a subject. In these same languages an unambiguous indication of the subject's number, person, and/or gender is generally found in the verb; e.g., Russian:

Otec uvidit mal'čika.

Father-NOM will-see-3sg boy-ACC

'Father will see the boy.'

In many other languages the verb phrase marks not only the subject but also other actants. Generally, such marking is sufficiently unambiguous to convey grammatical information about the grammatical subject even if the subject noun phrase is not present in the sentence; e.g., in Bulgarian:

Učiteljat šte daje knjigata na učenika.

the-teacher will give-3sg the-book to the-pupil

cf. with deletion of all noun phrases:

Šte mu ja daje.

will to-him-DAT her-ACC give-3sg

'He will give him it.'

It is important to note that in the above sentences there is no ambiguity as to which noun phrase or verb-phrase marker indicates the grammatical subject, which the grammatical direct object, and which the grammatical indirect object. This lack of ambiguity with respect to the subject is to be expected since the subject is in fact the organizing principle around which the syntax of the sentence is organized in these languages.

Now let us examine the situation in contemporary Georgian. As is well known, there is no general agreement among specialists in Georgian as to what is the grammatical subject. As can be seen from the chart below, practically all possibilities have been suggested:

Concepts of Subject

	Šanije (1953)		
	Čikobava (1967)	Vogt (1938)	Marr-Brière (OGeo)
	Aronson (1970)	Harris (1981)	(1931)
Present-Future	NOM	NOM	NOM
Aorist	ERG	ERG	NOM (?)
Perfect	DAT	NOM	NOM (?)

Chart 1

For the sake of simplicity, in the first part of this discussion we shall define the grammatical subject as that noun phrase which, when animate, triggers number agreement in the third person plural (see Aronson 1970:294).⁴ According to this definition, the grammatical subject of the various conjugations and screeves will be as follows:

Arguments:

Series- > Conjugation	Present, Future	Aorist	Perfect
I. ("Active")	NOM	ERG	DAT ⁵
II. ("Passive")	NOM	NOM	NOM
III. ("middle")	NOM	ERG	DAT
IV. ("indirect") ⁶	DAT	DAT	DAT

Chart 2

Operating with this definition of subject, let us see some of the problems that the concept of grammatical subject presents in Georgian. First of all, we should note that unlike most languages of Europe, Georgian possesses a case, the ERG, which has as its sole function to mark the grammatical subject of the sentence (in terms of the above definition).⁷ In this respect it differs from the NOM of languages of Europe which, as was noted above, has functions in addition to that of marking grammatical subject.

Among the major factors leading to difficulties in defining "subject" in Georgian are the well-known phenomenon of case variation for the same actant depending on screeve, widespread structured homonymy between forms marked for different voice categories (see Aronson 1985 for details), the fact that the valence of a verb (or sentence) very often cannot be recovered from the surface structure of the sentence, and the

growing lability of many verbal forms. The latter two phenomena will be discussed below.

Valence and the “unemphatic pronoun drop” paradox. Unemphatic pronoun drop (see Harris 1981:32-ff.) is found in many languages of Europe. It usually results in little or no loss of grammatical information, since the person, class, and/or number of the deleted pronoun is generally also marked by the verb form itself. An instructive example comes from modern Hebrew, where unemphatic pronoun drop is normal in the past and future tenses (where person/number is marked by the verb form) but not in the present (which marks gender/number, but not person):

[ani] e-šmor ‘[I] shall guard (fut., 1sg.)’
 [ani] šamar-ti ‘[I] guarded (past, 1sg.)’

but:

ani šomer (m.) / šomeret (f.) ‘I am guarding (present)’

In Georgian unemphatic pronoun drop works similarly when the deleted pronoun is first or second person. Since there is a meaningful trace of the person in the verb form, there is usually no loss of grammatical information here;⁸ compare: *me šen mas mogcem* and *mas mogcem*, both meaning ‘I (*me*) shall give it (*mas*) to you (*šen*, g-, indirect object) or *man me momkla* and *momkla*, both meaning ‘he (*man*) killed me (*me*, m- direct object). But with 3d person actants, the dropping of the pronoun can leave no trace of the presence of that actant or can leave an ambiguous trace. This can occur both with direct objects and indirect objects; examples of the former are:

- a. Mas çers. ‘He is writing it.’
- b. çers. ‘He is writing it.’ | ‘He writes, is writing.’

Not rare are examples of I. conj. verbs with preradical vowel *i-* (Tschenkéli’s T²) occurring intransitively; see examples below.

Examples with indirect objects are:

- a. with preradical vowel *a-*:
 - a. Açi siḡvdilis meṭi veraperi (mat) ga-a- šorebs. ‘Now nothing other than death can separate them. [After KEGL, s.v. gaašorebs.]
 - b. (Man) balaxs tesli gaašora. ‘He separated the seeds from the grass. [ibid.]

Although the verb form is identical in both sentences, in (a.) it is bipersonal while in (b.) it is tripersonal. Unemphatic pronoun drop will result in homonymy and potential ambiguity.

Note, too, that the preradical vowel *-a* functions to mark Tschenkéli's T¹ (transitive in neutral version, i.e., without indirect object) as well as T⁴ (superessive) and T⁵ (with direct and indirect object), and therefore does not give any information as to the presence or absence of 3d person indirect objects.

b. with preradical vowel *e-*:

a. Mat gadaego. 'He sacrificed himself for them.'

b. Gadaego. 'He sacrificed himself for them.' | 'He died.'
[Tschenkéli]

c. with preradical vowel *u-*:

a. Šemdegši V. Topuria mi-u-tit-eb-d-a, rom... 'Later V. Topuria indicated that...' [Gamqrelije and Mač'avariani, *Sonanṭta sistema...*, p. 017]

b. Šemdegši V. Topuria (mat) miutitebda, rom.... 'Later V. Topuria indicated to them that....'

In the above examples the (b.) sentence can be interpreted either with drop of a third person pronoun but otherwise identical in meaning to the corresponding (a.) sentence, or as not having an indirect object. Although instances where the marker *u-* (traditionally, the 3rd person objective version marker) does not indicate a DAT actant are rare, the remaining types of homonymy, namely instances with preradical vowels *a-* and *e-* are quite common.

d. before verbal root: Another source of ambiguity are the so-called 3d person indirect object markers. According to the norms of the literary language the markers and their distribution in immediate pre-root position are: *s-* before /t, ʈ, d, c, ɕ, j, č, ċ, ʝ /; *h-* before /p̥, k, k̥, g/ and facultatively (though rarely) before all other consonants, and *∅-* before all vowels and generally before all consonants not listed above. Clearly, in those instances when the object marker is *∅-* there can be ambiguity, as in the (b.) sentence below:

a. Ramdenime çeliğa mas darča? 'How many years then remained to her?'

b. Ramdenime çeliğa darča? 'How many years then remained to her.' | 'How many years remained?' [Mnatobi 1971, 47-9:109]

In actual usage there are numerous instances when a given verb form with an associated 3d person indirect object has no marker, despite the normative rule:

...kategoria..., romelic cidan čamovardnils gavs. '...a category which is similar to something that has fallen down from heaven.' (N. Dumbaje in Aronson 1982:444)

According to the literary norm, one would expect the form *hgavs* here. In other instances, an overt marker (*h-* or *s-*) is found, even though there are no 3d person direct objects associated with the verb:

Is pïesa, romelic am sami çlis çinat daşçera... 'The play that he wrote these three years ago...' [*Mnatobi* 1971,47-11:104.]

...amdeni adamianis ertianma xarxarma še-s-jra aremare. 'The united guffawing of so many people moved the surroundings.' [ibid.]

Mo-s-čanda ra? 'What was to be seen?' [ibid.]

In none of the above sentences is there a dative indirect object and according to the norm one would expect *daşçera*, *šejra*, *močanda*.

From the above, it should be clear that there are in Georgian numerous instances when one cannot determine the valency of a verb/sentence on the basis of the surface structure of the verb/sentence.

Lability for voice. By lability for voice we refer to those instances when one cannot determine the voice (transitivity/intransitivity) of a verb form or utterance.⁹ Connected with this is the difficulty or impossibility of determining subject/(direct) object relations. As can be seen from Chart 2 above, of the three cases that can mark arguments, only the ERG is unambiguous with respect to the argument marked. Now compare the following sentences:

- a. Masčavlebel-ma gaķvetil-i daiçqo.
teacher-ERG lesson-NOM he-began-it (aor)
The teacher began the lesson.
- b. Gaķvetili daiçqo.
lesson-NOM begin-aor
???
- c. Daiçqo.
begin-aor
???

In (a.) there is no ambiguity. The verb is transitive, I. conjugation, *masčavlebelma* is grammatical subject and *gaķvetili* is grammatical direct object. In (b.) and (c.), however, it is impossible to determine grammatical relations from the forms given alone. So, for example, (b.) can be glossed as:

[he] began the lesson (< Man [ERG]) *gaḵvetili daiçqo.*)

which is structurally identical with (a.) except for the dropping of the subject noun phrase in the ERG, or as:

the lesson began

which makes *daiçqo* then a II. conjugation verb, with *gaḵvetili* now the grammatical subject in the NOM and with no unemphatic pronoun drop. The same two analyses are valid for (c.), except that with the reading 'he began it,' both the subject (ERG) and direct object (NOM) pronouns are deleted and with the reading 'it began,' the subject (NOM) pronoun is dropped. How can we determine the syntactic structure of the ambiguous (b.) and (c.) and therefore determine what is the grammatical subject of the sentence?

One possible answer is to ask the question of what the verb form in question would be in the future or perfect screeves. Compare now (d.) and (e.):

- d. *Masçavlebel-i gaḵvetil-s daiçqebs.*
teacher-NOM lesson-DAT he-will-begin-it
The teacher will begin the lesson.
- e. *Gaḵvetil-i daiçqeba.*
lesson-NOM it-will-begin
The lesson will begin.

If (b.) or (c.) is transformed in the future screeve into (d.), then the form is I. conj., transitive, and with deletion of grammatical subject (b.) or subject and direct object (c.). If, on the other hand, (b.) or (c.) has as its future screeve transform (e.), then it is a II. conj. intransitive verb with deletion of the grammatical subject in (c.). To summarize:

	A. I. Conj.	B. II. Conj.
	'begins sthg.'	'sthg. begins'
FUTURE	<i>daiçqebs</i>	<i>daiçqeba</i>
PERFECT	<i>dauçqebia</i>	<i>daçqebula</i>
AORIST	<i>daiçqo</i>	<i>daiçqo</i>

Chart 3

In other words, the determination of the structure of *daiçqo* is dependent upon which paradigm it finds itself in: if it is in paradigm A, it is transitive, with ERG subject and if it is in paradigm B, it is intransitive with NOM subject. But this means that we must rely on a *paradigmatic* definition of subject, which should be basically a *syntagmatic* category.

Obviously this situation presents little difficulty for a grammar whose purpose is to generate Georgian utterances, i.e., to account for how the speaker encodes utterances. But it presents grave difficulties for a grammar that attempts to explain how the listener is able to decode the utterance, i.e., for a grammar that is based primarily on an analysis of surface features. From the latter point of view, one must conclude that a Georgian form such as *daiçqo* is just as labile with respect to voice as is, for example, English *began*, taken in isolation; cf. the following sentences:

- a. The teacher *began* the lesson
- b. The lesson *began*.

The Georgian verb in the above example offers as much information as to voice as does the English. Compare now equivalent Russian examples:

- a. Učitel' *načal* urok.
- b. Urok *načalsja*.

where there is a formal distinction between transitive and intransitive.

It should not be thought that the above example is an isolated one in Georgian.¹⁰ What is strange and unique about Georgian is the fact that despite the complex and rich verbal inflection of Georgian, a significant number of verbal forms turn out to be quite labile for voice (i.e., transitive/intransitive) relationships. Below, we shall attempt to list a number of examples of such lability in verbal forms.

Formally predictable lability. In the above example the lability of the form *daiçqo* is formally predictable: the aorist series screeves of I. conjugation verbs with the preradical vowel *i-* and the corresponding II. conjugation in *i-* aorist series screeves will be homonymous in all forms except for the 3d person plural. A number of examples of such resultant lability are given in Aronson 1985. Let one example suffice:

Ra čašuši daičira?! (J̄M, p. 75)

- a. What kind of a saboteur did he catch? (I. conj.)
- b. What kind of a saboteur was caught? (II. conj.)

Lexical lability. In a number of forms lability for voice is apparently not predictable on formal or semantic grounds, but appears to be a feature of a given lexical item. Such an example in a verb that is formally I. conjugation is *daagvianebs*, which can be used transitively, 'cause someone to be late' and intransitively 'be late'; e.g.:

- a. Davagvianet, vinaidan lekciebi gvian damtavrda.
'We were late because the lectures ended late.' (Aronson 1982:125)
- b. Amxanagis mosvlam teatrši damagviana.
'[My] comrade's arrival caused me to be late to the theater.'
(KEGL s.v. *daagvianebs*)

Again, the form *davagvianet* in (a.) by itself can be interpreted as meaning either "we were late," or "we caused him/them to be late." Note that the form in (b.), *damagviana*, is no longer labile because of the presense of the overt [direct] object marker, *-m-*.¹¹

Other examples:

- c. Menavem gainapira.
'The boatman landed.'
- d. Ertxel gainapira kidev moxuci Gurieli.
'He once again took old Gurieli aside.'
- e. Raeb's ambob? Xom ar gabodebs? [KEGL]
What kinds of things are you saying? Are you raving?
- f. ... — "..., sicxe tu gabodebt, Tvara, ras brjanebt ucnaurs?"
[KEGL]
'Is the heat making you rave, Tvara, what are you saying to the stranger?'

Intransitive I. conjugation verbs. Related to the lexically labile verbs are I. conjugation verbs that in no way differ formally from the overwhelming majority of verbs of that class, but which, unlike the majority, apparently occur only intransitively. There is a rather significant number of such verbs. The two most common patterns are verbs with the preradical vowel *i-* (traditionally often "subjective version") and verbs with the indirect object markers *mi-*, *gi-*, *u-*, etc. (traditionally "objective version"). Examples with so-called neutral version forms are also found, however, e.g.:

- čamožinžglavs (aor. čamožinžgla) 'drizzle (rain)'¹²
- asarsalebs (aor. asarsala) 'chatter'
- aabžuťebs (aor. aabžuťa) 'begin to twinkle, shimmer'

Examples with preradical vowel *i-*:

- moinaqrebs (aor. moinaqra) 'have a bite to eat'
- daircxens (aor. daircxina) 'become embarrassed'
- gaiocrebs (aor. gaioca) 'be astonished'

Examples with preradical vowel *mi-*, *gi-*, *u-*; these have an indirect object with no direct object.

unanavebs 'sing someone a lullaby'
 miusamjirebs 'give someone condolences'
 mousvenebs 'leave someone in peace'

Clearly, the widespread notion that in Georgian, unlike the situation in most European languages, there is a close correlation between form (conjugation class) and function (transitive/intransitive, active/passive, etc.) is not totally true: I. conjugation verbs can be transitive or intransitive as in the above examples; II. conjugation verbs can be active or passive; etc. And while it is important to show that the correlation between form and function in Georgian has been exaggerated, there is nothing typologically unusual about covert, rather than overt marking of transitivity in verb forms. As mentioned, this is the norm in French, German, Russian, etc. But the presence of covert marking for transitivity combined with unemphatic pronoun drop means that there is no necessary syntagmatic indication of transitivity in a given *sentence*. In simplest terms, within the confines of the sentence, the sentence *çers* can be interpreted as meaning either 'he is writing (in general, no specified object)' (< *Is çers*) or 'he is writing it' (< *Is mas çers*).

Connected with this is the question of "deletion of the direct object" in idioms and fixed expressions. With a number of I. conjugation verbs (usually relative) a "fixed" direct object can be and often is deleted. The question arises: in such instances are we still dealing with a transitive verb with unemphatic pronoun drop or with a verb that has either become intransitive or which is labile as to voice. Examples include:

- a. Martlac roṭmistrma daureḡa zari. [KEGL]
 The cavalry captain really rang his [DAT] bell [NOM].
- b. Amxanagma daureḡa.
 [His] friend called him [DAT].

In (b.) the direct object noun phrase was historically *teleponi* 'telephone,' i.e., the sentence was originally 'He rang his telephone.' But is the sentence in (b.) synchronically transitive or intransitive? And what criteria can we use to determine this?

Inversion constructions. Occurring with I. conjugation verbs, these are often examples of lexical lability. In these examples, the verb can be conjugated either "objectively," i.e., with inflection of the DAT noun phrase or "subjectively," with inflection of the NOM phrase; e.g.:

- a. *axvelebs*
he coughs

From the point of view of English, there is no problem in determining the meaning of this form: it is the equivalent of English 'he coughs.' But from the form in (a.) above there is no way of determining the syntax of the construction, i.e., there is no way of determining the case of the (deleted) subject. In fact, without unemphatic pronoun drop we could have the following possibilities:

- b. is *axvelebs*
he-NOM he-coughs
c. mas *axvelebs*
he-DAT he-coughs

According to our criteria, the verb in (b.) would belong formally to the I. conjugation (despite the fact that it is unambiguously intransitive) and the verb in (10c.) would belong to the IV. conjugation (see note 6), since the form meaning 'they cough' would have plural marking in the verb for the DAT argument: *mat a-xvel-eb-t*.¹³ Yet, with unemphatic pronoun drop, the third person form *axvelebs* gives us no information as to the case of the subject, as to the syntax of the construction. This can be determined, again, only paradigmatically:

I. Conjugation

v-a-xvel-eb	v-a-xvel-eb-t
a-xvel-eb	a-xvel-eb-t
a-xvel-eb-s	a-xvel-eb-en

IV. Conjugation

m-a-xvel-eb-s	gv-a-xvel-eb-s
g- a-xvel-eb-s	g-a-xvel-eb-t
a-xvel-eb-s	a-xvel-eb-t ¹⁴

If *axvelebs* is in the same paradigm as *vaxveleb*, then it is a NOM-ERG-DAT subject construction; if it is in the same paradigm as *maxvelebs*, then it is an inversion construction, with the subject always in the DAT. With unemphatic pronoun drop, the 3d person singular verb form is labile and the syntagmatic determination of what the subject is becomes almost impossible.

Lability in II. conjugation verbs. II. conjugation verbs frequently occur with active semantics (generally with an animate subject; with inanimate subjects the meaning is generally passive) (see Jorbenaje 1984:46-49) and in fact such forms at times can be used synonymously with transitive constructions. What is particularly interesting is the fact that the "transitive" use of II. conjugation verbs appears to be limited to the present series, where the formal opposition between the I. and II. conjugations is unambiguous. In the aorist series, where the semantic

opposition active/passive is at least clearer, the formal opposition can be lacking;¹⁵ e.g.:

Present series:

ıçers = he is writing [it] for himself | ıçereba = it is being written/he is writing it

Aorist series:

daıçera = he wrote it for himself | it was written

(Note also the following example, where one and the same II. conjugation form can occur with active and passive meaning and in both instances with an animate subject:

Passive: İbertqeba igi (viğacisgan). 'He is hit (by someone).'

Active: İbertqeba igi. 'He is hitting (i.e., he is beating the dust out of sthg.) (Suxiřvili 1981:193)

We are again faced with a problem of determining transitivity. If, as we have seen, verbs that are formally I. conjugation are not necessarily transitive, can we assume that II. conjugation verbs are necessarily intransitive? In the present and future series of screeves, as is well known, there is often no way of distinguishing between what are in the aorist series screeves direct and indirect objects:

Present/Future: Misçers İvanes çerils.

He'll write John (DAT) a letter (DAT).

Aorist: Misçera İvanes çerili.

He wrote John (DAT) a letter (NOM).

Present/Future: Levani çems das saçukars řehpirdeba.

Levan will promise my sister (DAT) a gift (DAT).

Aorist: Levani çems das saçukars řehpirda.

Levan promised my sister (DAT) a gift (DAT).

In the above examples, we can determine that one of the DAT forms in a present or future screeve is a direct object by transforming the construction into the aorist, where the opposition direct object/indirect object is disambiguated. But in the instance of forms such as:

Gazetebi ert ambavs ıçerebian. 'The newspapers are writing [about] one thing (DAT).'

there is no possibility of such a transformation; the construction does not exist in the aorist series. Further, *ert ambavs* does not pattern as an indirect object; verbs of the formal class of *ıçereba* mark indirect objects with the preradical vowel *e-*, i.e., *eçereba*.

A particularly common type of lability in Georgian is found in relative II-conjugation verbs in *e-*. Although most grammars would have you believe that the preradical vowel *e-* indicates the presence of a DAT (generally an indirect object), there is a significantly large number of forms which either (a.) never take a DAT or (b.) may or may not have DAT. As an example of the first type, consider the following:

Me *ḳi*, aravis ar gavaprtxileb, mexivit *davecemi tavze*. (J̄M 158)

[A soldier is describing how he will return home after the war:] 'I will not warn anyone; like a thunderstroke I will:

a. fall [on them].'

b. fall on my head.'

If we accept the first reading, then we are dealing with a relative passive with an indirect object, Tschenkéli's RP¹, *daecema mas tavs (tavze)* 'j-n/et. überfallen, über j-n/et. herfallen, sich auf j-n/et. stürzen.' But note that in this reading, the indirect object (*mas* or *mat*) has been dropped. If we accept the second reading, we have Tschenkéli's RP¹ *ohne i.O.* [without indirect object]: *daecema* 1. (auf et.) fallen/stürzen; 2. fallen (im Kampf) umkommen; 3. gestürzt w.; 4. fig. (moral.) sinken/herunterkommen. (In this reading, the postpositional phrase *tavze* has its literal meaning 'on head.') What is critical here is the fact that unemphatic pronoun drop in such instances leaves no trace of the deleted noun phrase in the sentence.¹⁶

The animacy hierarchy. In many languages opposing direct object and indirect objects, there is a clear tendency in sentences containing both for the direct object to be impersonal (or inanimate) and the indirect object personal (or animate). This is clearly true for Georgian, even more than for most other European languages; see Deeters 1930:97.

Examples of this can be seen in such constructions as the following:

Man mas daureḳa. 'He phoned her (indirect object).'

Man mas umṭera. 'He injured her (indirect object).'

Instructive in this respect is the contrast between *moismens* with direct object, translated by Tschenkéli as '(sich) et. anhören, et. hören, zuhören, horchen, lauschen; auf et. hören; et. erhören' and *mousmens* with indirect object, translated as 'j-m/e-r Sache zuhören, e-r Sache lauschen, j-n/et. anhören.' In the first instance the object is given only as inanimate; in the second, it can be animate as well as inanimate.

In contemporary Georgian, particularly in II. conjugation verbs there seems to be a strong tendency towards an animacy hierarchy, which operates in the following manner in the present, future, and aorist series:

1. Given two non-agentive actants, one of which is animate and the other inanimate, the animate will be DAT and the inanimate DAT/NOM; 3rd person number agreement if present will be with the DAT.

2. In bipersonal non-I. conjugation verbs, given two actants, number agreement in the third person will be with the animate, as opposed to inanimate actant.

3. In I. conjugation verbs, given an inanimate NOM-ERG and an animate DAT-NOM, there is a tendency for the animate DAT-NOM to show 3d person number agreement.

In general, number agreement tends, then, to be correlated with animacy, a nominal (inherent, covert) category, rather than with syntactic categories, such as subject, direct object, indirect object. Much evidence in favor of this has been given by Jorbenaje (1984) and Kiziria (1985). We shall now attempt to document this tendency; note the following sentence:

Itxovdnen mas qvelgan, undodat misi naxva, *ainteresebdat* rogor iqo is, magram tito-orola adamianis meṭi aravin ar icnobda. 'They would ask for him everywhere, they wanted to see him, how he was interested them, but aside from one or two people, no one knew him.' (*Mnatobi* 1971, 47-11:105.)

Here the "subject" (NOM) is inanimate,¹⁷ 'the fact of how he was,' while the "direct object" (DAT) is animate, 'them,' i.e., 'the people.' The number agreement is with the DAT, and thus according to my earlier definition of subject (Aronson 1979:294), the *subject* of this I. conjugation form should be in the DAT. But a problem arises: Should the DAT in this subject have been singular, i.e., *ainteresebda rogor iqo is*, 'she was interested in how he was,' there, again, is no syntagmatic way of determining what the subject is.

A real problem arises in verbs which normally take both an animate subject and object. Such, for example, is the relative II. conjugation verb *še(h)xvdeba* 'someone meets, encounters someone.' This verb apparently can occur with either the DAT or the NOM as "subject" (again defined in terms of third person number agreement). In the following example the subject is NOM:

Teaṭrši nacbobs ševxvdi. 'Ja vstretil znakomogo v teatre.' 'I met an acquaintance (DAT) at the theater.' (RKL, s.v. *vstrétit'*)

Such a construction is in Tschenkéli's terms RP⁵, a relative passive with the subject in the NOM. But Tschenkéli also lists a RP⁵(OR) (= objektive Reihe, subject in DAT), examples of which abound. In many of these

examples, as is to be expected, the NOM is inanimate and the DAT is animate:

Tu *ḡargi dari daudgat*, *prinvelebi mgzavrobas ḡkara ataveben...*, tu *cudi dari ṣexvdat*, *ugviandebat*. 'If they (DAT; pl) encounter good weather, the birds (NOM; pl.) end their trip quickly..., if they (DAT; pl.) encounter bad weather (NOM), they (DAT; pl) are late (lit. it (= the trip, NOM) is delayed to them (DAT)).' (KEGL, s.v. *ṣexvdeba*.)

The above example should present no difficulties if we have a hierarchy in which number agreement between NOM and DAT is not determined by case, but rather by animacy. With the verb forms *daudgat*, *ṣexvdat*, *ugviandebat* the agreement is with the DAT, which is animate, and not with the NOM, which is inanimate. With *ataveben*, the NOM is animate and the DAT is inanimate; agreement is with the animate NOM.

But problems arise when both arguments are animate, e.g., in the following examples:

Ḳoḡḡcaant baḡḡebi raḡac cxvirs aceminebdnen, *izmorebodnen da kalebma ipikres*, *albat baṡonebi tu ṣexvdato*. 'The *Ḳoḡḡcajes*' children were sneezing, stretching and the women thought that they (DAT; pl.) probably might meet some gentlemen (NOM; verb form is sing.).'

Dacinvit ṣegvxvdnen. 'My *vstretili ix radušno*.' 'We (DAT) met them (NOM) with laughter.'¹⁸ (RKL, s.v. *vstřétit*)

Is xalxi, ḡḡan rom ṣegvxvda, *ṣejlebs amas*. 'Those people whom we (DAT) encountered a little while ago are capable of doing this.' (*Mnatobi* 1971, 47-9:109.)

In this sentence it is clear from context that the translation is as above, and not 'who encountered us.'¹⁹

From the above, we can see that there is a tendency toward the hierarchization of the actants in the clause according to animacy/inanimacy. This hierarchy is in conflict with the case hierarchy (for I. and III. conjugation verbs in the present, future and aorist series — ERG-NOM-DAT). As a consequence, there are numerous instances when a given clause does not provide information sufficient to allow determination of what the subject is. Neither the animacy hierarchy nor the case hierarchy, however, are sufficient to account for instances when both actants are animate and the number agreement can be with either NOM or DAT.²⁰

Conclusions. To summarize, in Georgian, in most instances, information from the verb is sufficient to give information as to whether a

construction is transitive or intransitive or what the grammatical subject of the sentence is (using the definition described above) either by itself or in conjunction with information from noun phrases (case). But there is a significantly large number of instances when such a determination appears impossible based solely on information in the given sentence or clause. The question naturally arises: are we in fact dealing with a nominative (or ergative) system here, or with one at some (probably early) stage of transition from a transitivity-based sentence structure to one not transitivity based? The Georgian structure, although there are some instances where an animacy hierarchy appears to play a role, does not resemble the system found in active languages as described by Klimov, nor does it resemble the system typical of English, both systems lacking to a greater or lesser extent the grammatical opposition of transitive/intransitive.

What is unique about Georgian, it seems to me, is the fact that a highly complex and developed morphological system results in such a high degree of lability or neutralization of formal oppositions (see Aronson 1985).

The question remains, how do Georgians interpret sentences which due to [surface?] homonymy and unemphatic pronoun drop appear to contain insufficient information for determination of voice relationships, of subject/object relationships, and even as to whether NOM, DAT, or ERG actants are present in the sentence?

I do not have the answer to this question. But I would like to suggest that previous attempts at a definition of grammatical subject for Georgian are either 'unmotivated, arbitrary, or else do not always allow for predictability on the level of the actual utterance without recourse to a paradigmatic level or to an underlying structure. (It follows, of course, that if the definitions of subject are not satisfactory, the same is true for definitions of objects, both direct and indirect.)

One might claim that the necessary information is to be found *outside* of the sentence, in the greater context. This may very well be true, but if it is, it must be noted that this is typologically quite different from the systems that characterize both nominative and ergative languages, and is therefore of great significance.

I have difficulties in parsing (and hence in interpreting and understanding) Georgian sentences; Georgians, apparently, 'do not.'²¹ That the Georgians can understand such sentences without difficulty is clear proof that there is present on the surface level in Georgian sufficient information to allow Georgians to decode the overwhelming majority of utterances in their language without difficulty. In lieu of a solution, I should like to suggest that to solve the problem of the subject

(and concomitantly, of voice relationships) in Georgian it may well be necessary to look at the data from a fresh point of view, a point of view that does not presuppose the universality of transitive/intransitive nor the universality of the syntactic category of grammatical subject. I should like to further suggest that the notion "grammatical subject" is only marginally applicable to literary Georgian. Rather than serving as the starting point for the description of the syntax of the language, the notion of subject in Georgian becomes the debated (and debatable) result of complex analysis. The very fact that there is so much debate as to which noun phrase in a given construction should be regarded as grammatical subject is a further argument against the general applicability of such a category in Georgian.

If we accept this and the weakness of the opposition transitive/intransitive in Georgian, a very significant consequence follows: typologically Georgian deviates significantly from both the nominative and ergative structures. I earlier characterized Georgian as a basically nominative language (Aronson 1970:296-97), but that analysis was based almost totally on paradigmatic criteria. When we look at the syntagmatic data, such a conclusion becomes much harder to justify.

In analyzing Georgian, the question that must be asked is not "what preexistent typological mold does Georgian fit into?" but rather "what information is available on the level of surface structure in Georgian and how do speakers of that language utilize it to decode the messages that they hear?" And I hope to have demonstrated, at least in part, the importance of a grammatical analysis that attempts to account for how addressees are able to decode the messages that they receive, relying totally on surface information.²²

The University of Chicago

Notes

¹Stankiewicz's definition of inflection raises serious problems in Georgian and may not be applicable here: in a significant number of verbs in Georgian the presence of a non-past screeve does *not* imply the presence of past, future, or perfect series screeves (e.g., the forms listed in Tschenkéli *GDW* as having only present series screeves). Yet, as a general rule, the presence of aorist series screeves implies the existence of present and future series screeves, i.e., the relationship is one characterized by Stankiewicz as typifying derivation.

²So, for example, Klimov (1972) has shown that such syntactic categories as direct object and indirect object are restricted to languages of the nominative (and for him,

ergative) structures. For a claim that English does not distinguish between direct and indirect objects, see Aronson (1977:210).

³In languages such as Russian and Latin, even when there is no overt NOM noun phrase that marks the subject of an “impersonal” sentence, there is still generally a grammatical marking of person-number or gender number, which can be said to mark a grammatical subject; e.g.:

Latin: *pluvit* (3sg) ‘it is raining

Russian: *mne ne spitsja*

to-me-DAT not sleep-3sg-intrans. ‘I don’t feel like sleeping.’

⁴Yet this definition presents serious problems. Consider, for example, the following sentence (Gogebašvili, cited in Kiziria 1985:109): *Šimšilma jlier šeaçuxat lekvebi* ‘Hunger caused the puppies great pain,’ where despite the presence of an (inanimate) ERG there is 3d person NOM number agreement. Should we regard the ERG as subject or should we view the NOM as subject because it shows number agreement? The increase in NOM number agreement can only lead to a weakening of the force of our subject-definition rule. For further details, see Kiziria 1985.

⁵As Harris (1981:302) has pointed out, this definition results in a shift of subject within the perfect series screeves in the normative language. In the spoken language there is a tendency to have the number agreement be consistently with the DAT in both I. and III. conjugation verbs.

⁶For the purposes of this paper, IV. conjugation includes not only those verb forms which have “active” forms in the present series and relative passive (in *e-*) forms in the future and aorist series (e.g., pres. *miqvars* ‘I love her,’ but fut. *meqvareba*,) but all verbs, be they I., II, III, or IV. conjugation, which show 3d person plural number agreement with a DAT argument in the present, future, and aorist series. These are the verbs that are characterized by Tschenkéli *GDW* as O[bjective]-R[eihe], e.g., *uçirît guli* ‘they (DAT) are extremely sad (III. conj.)’, cf. *auçirdebat* ‘they (DAT) will become extremely sad (II. conj.)’.

⁷The subject marking strength of the ERG is so great that it can be used even when the verb (of speaking) is omitted, e.g.:

Manana-m čemi çigni momecio.

Manana-ERG my book-NOM (d.o.) give-me-it-indirect speech particle

Manana [said] that you should give her her book. | Manana said, “Give me my book!” (Aronson 1982:212)

In this example, the verb *tkva* ‘she-said-it’ is omitted, yet the presence of the ERG is sufficient to indicate the deletion of a verb which takes a 3d person ERG as its subject.

⁸We shall not discuss here possible ambiguities arising from the loss of desinential *-s* before the plural marker *-t* or the constraint against first person *v-* immediately preceding the second person marker *g-*. For examples of these, see Tschenkéli 1958:352-55, 360-64.

⁹Jorbenaje 1983 (passim.) gives a great variety of examples of lability for voice and valence.

¹⁰A Russian example, such as the stylistically marked

Mat' doč' ljubit.

mother-NOM or ACC daughter-NOM or ACC she-loves

'Mother loves [her] daughter.' or 'Daughter loves [her] mother.'

is in Russian relatively isolated and represents a formal neutralization of the NOM/ACC opposition. But the Georgian form *daiçqo* cannot be said to be a formal neutralization of a single opposition as in the Russian example. Note the "neutralizations" that would be involved in this example:

	<i>daiçqo</i> ^I .	<i>daiçqo</i> ^{II} .
VOICE	transitive	intransitive
SUBJECT	ERG 3sg	NOM 3(sg.)
DIRECT OBJ.	NOM 3d	— —

And, as we shall attempt to show below, the Georgian example is not isolated.

¹¹This gives us another method of disambiguating labile forms; if the form in question is part of a paradigm marking a direct object in the first and second persons, then the verb is transitive: *damagviana*, *dagagviana*, *daagviana*, etc. However, this is again a paradigmatic solution to a syntagmatic problem. Further, it presupposes that we can distinguish between *m-*, *g-*, *gv-* marking direct objects and the same formants marking indirect objects, which is not always possible.

¹²Although Tschenkéli (GDW) treats this as a middle verb, the formation is clearly not III. conjugation, but rather I. conj. See also *axvelebs*, discussed below.

¹³Some evidence that the construction *mas axvelebs/mat axvelebt* should be treated formally as IV. conjugation and not as I. conjugation is given by the perfect series screeves of this verb as given by Tschenkéli GDW: perfect *da-u-xvel-eb-i-a*. This is, to be sure, the same form as the perfect of the I. conjugation is *axvelebs*, but (if, in fact, *mas axvelebs* does have a perfect series) these two forms must be regarded as having different derivational histories:

	I. Conj.	IV. Conj.
Present:	v-a-xvel-eb	m-a-xvel-eb-s
Perfect:	da-mi-xvel-eb-i-a	da-mi-xvel-eb-i-a

As can be seen from the above, the I. conjugation verb shows the inversion associated with the formation of the perfect series of members of this form class, i.e., the NOM subject of the present and future series is converted by inversion into a DAT subject in the perfect. But in the case of the IV. conjugation verb, where inversion already occurs in the present and future series, there is no "tvisation," i.e., a DAT is not replaced by a postpositional phrase with *-tvis* 'to, for.' We do not get the expected **da-u-xvel-eb-i-a čem-tvis*. This seems to serve as evidence that the DAT in *mas axvelebs* is an underlying DAT subject and not indirect object, since tvisation apparently applies only to indirect objects and not to subject DATs.

¹⁴Note that the form *axvelebt* can therefore mean either 'you all cough (I. conj.)' or 'they cough (IV. conj.).'

¹⁵Suxișvili points out that there now exist active ("deponent") II. conjugation verbs in *i-* that have aorist series screeves that are formally I. conjugation, in that they can

take the ergative (1981:188-91). He views such forms as *iğimeba igi | gaiğima man* 'smile' as now entering into one paradigm, although historically they come from different conjugations. Similarly, the older relationship *is ixedavs — man gaixeda* 'see' (both forms I. conjugation) has been replaced by *is ixedeba — man gaixeda* (II conj. alternating with I. conj.) See also Jorbenaje 1984:49-50.

¹⁶It must be noted that this ambiguity is found only with a third person DAT. There is disambiguation in the perfect series, where the RP¹ will have the relative form *davcemivar tavze* and the RP¹ (ohne i.O.) will have the absolute form *davcemulvar*.

¹⁷We use the terms "animate" and "inanimate" rather than "personal" and "nonpersonal" on the basis of examples showing 3d person DAT agreement with nouns denoting nonpersonal but animate beings; see the example below where DAT plural marking is triggered by *prinvelebs* 'birds.' A similar example is:

Tu didi luḡma šexvdebodat [çiçilebs, ḡruxi] daanaçilebda, rom çaqlapva šesjlebodat mis švilebs. 'If the chicks (DAT) would encounter a big piece of bread (NOM), the hen would divide it up so that her children could swallow it.'

It should be clear that the opposition discussed here is, in Whorf's terms, covert rather than overt. It further is an innovation and not a relic of an earlier stage of Kartvelian (Ḳiziria 1985:106). It is typologically totally different in form and function from the class systems of the North Caucasian languages and should not be confused with the latter, nor should it be adduced in support of a putative earlier class structure for Kartvelian. (See Oniani 1985 for a convincing argument against the latter.)

¹⁸This example presents yet another problem. Since semantically this verb must be Tschenkeli's RP5(OR), it should show 3d person DAT number agreement. But then the marking of number agreement for a third person NOM becomes anomalous; i.e., we would expect *šegvxdā*. It is possible that there is a person hierarchy operating here in addition to the animacy hierarchy.

¹⁹In another sentence from the same short story the relationship between the two actants is also clear from context:

Ḳrimḡkis eloda, Soxumḡši unda šexvedroda, magram ver šexvda! 'She was waiting for Krimsky; he was supposed to meet her in Sukhumi, but he couldn't meet her!' (*Mnatobi* 1971, 47-9:109.)

It is most likely that the case marking is: 'he (NOM) was supposed to meet her (DAT) in Sukhumi, but he (NOM) couldn't meet her (DAT)!' A reverse reading is unlikely, since it would convey a nonvolitional meaning ('bump into') that does not combine readily with *unda* 'must' or *ver* 'cannot.'

²⁰In addition to the animacy hierarchy, there is also a person hierarchy, which excludes the possibility of 3d person number agreement when the subject is 1st or 2d person (Ḳiziria 1985:111).

²¹But see Šaniḡe 1953:454-5 (cited in Aronson 1984:36, n.9).

²²I have learned much about the structure of Georgian from my students Kevin Tuite and F. Gregory Paludis. Discussions with them helped greatly to crystalize my

notions on subject in Georgian and I would like to express my gratitude to them. It goes without saying that I assume responsibility for the ideas expressed here.

References

- Aronson, Howard I. 1970. "Toward a semantic analysis of case and subject in Georgian." *Lingua* 25:291-301.
- _____. 1977. "English as an active language." *Lingua* 41:201-16.
- _____. 1982. *Georgian, a reading grammar*. Columbus, Ohio.
- _____. 1984. "On homonymy in the Georgian verbal system." *Folia Slavica* 7:21-37.
- Čikobava, A. S. 1967. "Gruzinskij jazyk." in: Bokarev, E. A. and K. V. Lomtadze, ed., *Iberijsko-kavkazskie jazyki. Jazyki narodov SSSR*, 4. 22-61.
- Deeters, Gerhard. 1930. *Das khartwelische Verbum: vergleichende Darstellung des Verbalbaus der südkaukasischen Sprachen*. Leipzig.
- GDW. = Tschenkéli. 1965-74.
- Harris, Alice C. 1981. *Georgian syntax: a study in relational grammar*. Cambridge.
- ŽM. = Žgenti, Suliko. 1977. *Jariskacis mama*. Tbilisi.
- Žorbenaje, B. 1983. *Zmnis xmovanprepiksuli čarmoeba kartulši*. Tbilisi.
- _____. (= Džorbenadze, B.A.) 1984. "K voprosu o semantičeskoj xarakteristike form stradatel'nogo zaloga v gruzinskom jazyke." *Iberul-čavkasiuri enatmecnierebis čeličdeuli* 11:46-54.
- KEGL. = Čikobava, Arn., gen. ed. 1950-64. *Kartuli enis ganmartebiti leksikoni*, 1-8. Tbilisi.
- Kiziria, A. 1985. "Obiektis mier zmnis šetanxmeba mravlobit ricxvši tanamedrove kartulši." *Iberul-čavkasiuri enatmecniereba*, 24:100-12.
- Klimov, G. A. 1972. "K xarakteristike jazykov aktivnogo stroja." *Voprosy jazykoznanija* 4, 3-13.
- Marr, N. and Maurice Brière. 1931. *La langue géorgienne*. Paris.
- Oniani, A. L. 1985. "O grammatičeskoj kategorii klassa v kartvel'skix jazykov." *Voprosy jazykoznanija* No. 3-67-70.
- RKL. = 1956-59. *Rusul-kartuli leksikoni*, 1-3. Tbilisi.
- Šanje, Akači. 1953. *Kartuli gramatikis sapujvlebi, I. Morpologia*. Tbilisi.
- Stankiewicz, Edward. 1962. "The interdependence of paradigmatic and derivational patterns." *Word* 18:1-22.
- Suxišvili, Murman. 1981. "'Anuča ilanjga da igina' tu ilanjga da igina'?" *Kartuli sičqvis kułturis sačitxebi* 4:183-97.
- Tschenkéli, Kita. 1958. *Einführung in die georgische Sprache*, 1. Zürich.
- _____. 1965-74. *Georgisch-deutsches Wörterbuch*, 1-3. Zürich.
- Vogt, Hans. 1938. "Esquisse d'une grammaire du géorgien moderne." *NTS* 9:5-114, 10:5-188.

On Iranian Influences in Old Georgian

Roland Bielmeier

Iranian languages have influenced the Armenian as well as the Georgian language especially in their vocabularies. Most of these loanwords originate in the western group of the Middle or New Iranian languages: in Parthian, Middle Persian, and Early New Persian. In Georgian the situation is more complicated than in Armenian because a great part of the loans passed through (Pre-)Armenian into Georgian. But for some Iranian loans in Georgian we lack the correspondence in Armenian, cf. e.g., Geo. *kuča* "street" from early NP *kuča* < *kuyča* (Steingass 1892), diminutive to NP *kuy* < Zor. Pahl. *kwd* [kōy] "street, lane" (MacKenzie 1971:52). Thus the question arises, as to whether these Iranian loans in Georgian were transmitted through Armenian, and whether the Armenian word has since been lost, or whether they came directly from Iranian into Georgian. In some cases the shape of the Georgian form points to Armenian, or in other cases excludes transmission through Armenian; cf., e.g., Old Geo. *ambori* "kiss" where -r- points to Armenian transmission, cf. Arm *hamboyr* < Parth. **hambōd* (Manich. Parth. *'mbwy'd* [ambōyād] "kissed"), etc. (Schmitt 1983:87), but e.g., Old Geo. *pasuxi* "answer" < late MP, cf. Manich. Parth. MP *pswx* "id." (Boyce 1977:74) or early NP *pāsux* "id." vs. Arm. *patasxani* "id." (Hübschmann 1897:222), etc.

Besides this influence from West Iranian languages we also find traces of East Iranian influence in Armenian and Georgian. Henning (1958:93 and 1937:85) claimed that in certain cases in Armenian loanwords there was an East Iranian Parni element introduced through the Parthian language, that is "eastern Iranian forms brought by the Parni nobility of the Parthians to western Iran and to Armenia" (Frye 1969:158), cf. e.g., Arm. *margarē* "prophet" corresponding to Sogd. *m'rkr'y* /mārkarē/ "sorcerer" as against Parth. *m'rygr* /māregar/ "sorcerer, incantator" (Henning 1937:85 and 1958:93; see also Schmitt 1983:85 and Bailey 1985:106). "This view would accord with the history of the formation of the Parthians and their expansion over Iran." (Frye, 1969:158; see Schmitt 1983:83ff.). In Georgian, however, we find East Iranian loans

from the north, from Sarmatian or Alanic dialects, which belong to the East Iranian languages.

As far as Old Georgian is concerned, we depend on Old Georgian sources, of which the Old Georgian chronicle *Mokcevey Kartlisay* (MK) "The Conversion of Georgia" is one of the oldest in existence and which has not been translated from any other language. It represents a history of Georgia till the ninth century and the composition of its oldest sections dates back to the seventh century. This part of early Georgian history is also treated in the "Georgian Annals" *Kartlis Cxovreba* (KCx), the oldest sections of which, on the history of the Georgian kings, were compiled by Leonti Mroveli in the 11th-12th century (see Gippert 1984:38).

For the investigation of Iranian loans in Old Georgian the most interesting part of the MK chronicle is the first chapter which describes Georgia's pre-Christian period and its conversion to Christianity. A list of the names of thirty-four pre-Christian rulers is given in this chapter. Insofar as these names are etymologically explainable, they are of Iranian origin. And I think it can be shown that some of these names are of Parthian or Middle Persian or New Persian origin, but some of them may well be of Sarmatian or Alanic origin. In addition, some of the Old Georgian names which occur in this chronicle or in other Old Georgian inscriptions or texts correspond to Iranian names in the Greek inscriptions on the northern coast of the Black Sea. But this does not mean automatically that those names are of East Iranian origin, because we also find names of Middle West Iranian origin in these inscriptions.

The succession of the kings in the two chronicles is sometimes unclear, e.g., *Saurmag* as the third king in MK. The name belongs with *Sauromaces*, of whom we know (see below) that he reigned in the fourth century, etc. We therefore cannot rely on it, and I, of course, argue only for an etymological connection of the names, not of the personalities.

The name of the second king in the MK list is *Parnavaz*, rendered by the classical authors as Φαρνάβαζος. Justi 1895:92 analysed the name as **farnah-bažu-* and translated "dessen Arm (Gewalt) vom Glück begünstigt ist." But Marr 1902:5 derived Geo. *Parnavaz* from an older **Parnavazd* in analogy to Geo. *Armaz-i* from **Armazd* which is generally followed nowadays (see Andronikašvili 1966:496ff.), notwithstanding the fact that the semantics of the second element in **Parna-vazd* remains somewhat unclear. This second element is doubtless present in late Avestan *Ašauuazdah-*, whose western correspondence is preserved in Greek Ἀρταβάσδος, Lat. *Artavasdes*, Arm. *Artavazd*. Mayrhofer 1979:I/24 tentatively suggests "vielleicht im Bereich von 'Kraft, Ausdauer'." The fact that there is no trace of final *-d* in Armenian *Parnavaz* has already been explained by Hübschmann 1897:89, where he

pointed out that this name was borrowed from Georgian, in which it occurs as the name of a Georgian king. In Old Georgian final *-zd* is simplified to *-z*, as can be seen from the famous placename *Armazi*, being related to Old Perisan *Auramazdā*, etc., which is also kept in Armenian as *Aramazd*. For an analogous development in Sogdian see Gershevitch 1961:§266.

Frye 1969:84f. (= 1976:155f.) considered the Georgian name an early borrowing from Old Median, not yet showing the later development *-rn-* > *-rr-* > *-r-* of Middle West Iranian, see, however, Schmitt 1983:78. In addition to this, instead of taking it as an Old Median borrowing, we could also think of Middle East Iranian mediation, cf. Sogd. *prn*, Oss. *farn* “glory.”

The next name in the list is *Saurmag* which is clearly of East Iranian origin. Through Ammianus Marcellinus (27, 12, 4), whom we date to the end of the fourth century, we know that the Romans appointed a king in Iberia whose name Ammianus wrote *Sauromaces*, and who was expelled again by the old Sasanian king Shapur II (309-379). This name may be connected with the ethnicon of the Σαυρομάται, to which the name Σαυρομάτης of some Bosporian kings (1st-3rd centuries; see Justi 1895:292f.) also belongs. The latter also occurs in the Black Sea names (see Zgusta 1955:§767), but its derivation is still disputed. Andreas and other scholars connected it with Old Indian *roman-* n. “Körperhaar bei Tier und Mensch” and took it to be a plural form, “the ones with black hair.” Others, such as Marquart and Bailey, connected it with the *Sairima*-tribal name in the Avesta (see Lommel 1926:153, Bailey 1945:26, etc.). Justi 1895:292 analysed the Latin form correctly as “gekürzt aus Sauro-m(ates) mit Affix aka.” To this corresponds Kurylowicz’s “hypocoristic root.” For instance, *Humāyā* is shortened to its hypocoristic root *Hum-* to which the hypocoristic suffix, in this case *-iča*, is added > *Hum-iča*, see Mayrhofer 1973:284.

Andronikašvili 1966:506f. connects the name of the twelfth king in the list, who reigned in Armazi, *Karzam* (Takaišvili misreads *Karram*, see Pätsch 1978:288, 291, n.18; *Kartam* in KCx 45 is a later corruption), with the Black Sea names KAPZEIΣ and KAPZ(Y)AZOΣ (Olbia, second century, see Zgusta 1955:§§126, 590). Following Miller 1886:257 and 1913:84f., these names are usually related to Oss. *karz* “harsh.” In my opinion, however, it is more appropriate to take the Georgian name as an East Iranian Bahuvrīhi compound consisting of *karž-* “miracle,” and *ama-* “power,” i.e., “possessing miraculous power,” cf. Buddhist Sogd. *krz* /*karž*/, Chr. Sogd. *qrž* “miracle” (Benveniste 1929:100; Gershevitch 1961:§§263, 355; for a final identification of Sogd. /*karž*/ and Oss. *karz* see Szemerényi 1970:424f.). I also see a connection between the first element

of Old Georgian *Karṣ-am* and the above-mentioned Black Sea names, and a connection between the second element of KAPZ-(Y)AZOΣ and that of the above-mentioned Geo. *Parna-vaz* < **Parna-vazd* or Arm. *Arta-vazd*, etc.

On the eighteenth king we find the following sentence: *da mepobda Ḡadami da dge k* (= 20). Takaišvili 1900:15, n.2 translates: "I carstvovaḡ Ḡadam i dneḡ 20." The German translation by Pätsch 1978:292 renders it as "Und König war Ḡadami und zwanzig Tage..." In KCx 54/7ff. the name is corrupted to *Adami* and to *Ḡadana* in KCx 54/11 "name of the mother of *Adami*" respectively (see Androniḡašvili 1966:508). The syntax of the sentence is unusual and the second *da* does not make much sense. Unfortunately I have no access to the manuscripts but it seems to me that we have to understand this second *da* as a part of the name and to restore the name as *Ḡadami[z]da* or *Ḡadami[s]da*. Then the syntax is correct and the sentence means simply "and *Ḡadami[z]da* reigned 20 days" (according to KCx 54 he reigned three years). It is interesting that the father of *Ḡadami[z]da* is *Parsman Kveli* in KCx 54, who reigned together with *Mirdaṡ* and who was followed by his son *Ḡadami[z]da* as king. This fits perfectly with the succession in MK. There *Parsman* is followed by *Mirdaṡ* and *Ḡadami[z]da*. It also fits with the history of *Radamistus*, son of the Iberian king *Pharasmanes* in the first century who appointed first his brother *Mithridates* as king in Armenia who was later murdered by his nephew *Radamistus*, as we know through Tacitus (Tac. ann. XII:44ff.; for further details see Bielmeier 1990:42 and also Gippert 1984:39 with n.25). Moreover, the name can further be connected with several names from the Black Sea coast, of which ΠΑΔΑΜΕΙΣΤΟΣ (*Panticapaeum* third century A.D., *Tanais* second century A.D.) is the closest form to *Radamistus*, as Justi 1895:257 had already seen (cf. Zgusta 1955:§§186-191). Others are ΠΑΔΑΜΑΣΕΩΣ (genitive, Zgusta 1955:§186), ΠΑΔΑΜ-Ο-ΦΟΥΠΤΟΣ (Zgusta 1955:§188), ΠΑΔΑΜΣΑΔΙ(Ο)Σ (Bosporian king, fourth century, see Justi 1895:257). In the etymological interpretation of their first part *Radam(a)*- I follow Justi and Zgusta against Miller 1886:257. They understand it as an East Iranian development < **fratama*- "primus." For the second part Justi and Miller gave no explanation. Zgusta 1955:§187 following Vasmer 1923:48 related it to Av. *miṣda*- "Lohn, Gewinn, Vorteil" and interpreted "ersten Lohn habend." This is not very satisfactory. I see in the second element a secondary hybrid Middle West Iranian superlative formation (independently also Weber 1986:172, n.10), cf. Parth. *yšt* (Skalmowski 1967:86), MP *-ist* (Rastorgueva 1966:53). Georgian initial *-ṡ* corresponding to East Iranian *r*- remains difficult.

The early kings and also the *piṭiaxši*, the viceroys of Georgia, bore mostly Iranian names which were later more and more replaced by Christian Greek names. On the other hand the priests and the Katholikoi had Greek names and very rarely Iranian ones. But it is interesting that in the chronicle, the sister of Saint Nino was called *Šrošana*. This name goes back to Middle Persian *srōš* "the god Obedience" with the patronymic suffix *-ān*, descended from Av. *sraoša-m*. "Gehör, Gehorsam" and is also the name of a deity (see Benveniste 1945, Kreyenbroek 1985).

I have cited only a few examples of Iranian borrowings in Old Georgian. For further examples see Bielmeier 1985 and 1990, and also Gippert 1984. I would prefer to discuss a short passage from the first chapter of the MK chronicle. The passage has long been misread and misunderstood and has to be corrected on the basis of Iranian influence.

In line 19-22 of the first chapter of the MK chronicle, edited in Tbilisi 1979 by Gigineišvili and Giunašvili (p. 320), we read:

*xolo Sarkinesa kalaksa ebrzola atertmeṭ ttue da dadga Sarkinesa dasavalit
kerzo da daasxa venaqi da ruj (var. roj) gamoiṣo Ksnit, da dasxna kacni
meruveni daṣtagita rujsajta da hrkwian adgilsa mas Nastagisi.*

The English translation runs as follows: "But he [Alexander] fought the town of *Sarkine* eleven months and stopped west of *Sarkine*, laid out a vineyard and drew a canal (= *ruj/roj*) from the [river] *Ksani* and appointed people as canal overseers (= *me-ru(v)-e-n-i*) *daṣtagita rujsajta*; and this place has been called *Nastagisi*."

In this passage *daṣtagita rujsajta* remains unclear. The entry in the Georgian Annals relating to these events is much more condensed, and there is no trace of the untranslated words above which concern us here.

The famous Georgian lexicographer Sulxan-Saba Orbeliani (1658-1725) misinterpreted the passage. He read *da ṣtagita rusata* and translated *ṣtagi* as *rus satave*, i.e., "source of the canal" (Orbeliani 1966:112). In this he was followed by Čubinašvili (see Pätsch 1975:291, n.10) and Takaišvili. The chronicle was first edited by Takaišvili in 1890 (see Pätsch 1975:288f.) and was translated by him in 1900. The passage in question we find in Takaišvili 1900:6f. with note 2:

...postavil ljudej smotret' za kanalom; i ot «stagi» kanala nazvali èto mesto Nastagisi.

([Alexander] posted people to look after the canal; and because of the "stagi" of the canal, they called this place Nastagisi.)

In note 2 Takaišvili, following Orbeliani, translated *ṣtagi* as *načalo kanala*. He also referred to an article by Marr written in 1896, where the

latter cited a "fictitious" geographic term, *rotastak*, in the Armenian history of Agathangelos. In the new edition by R. W. Thompson (1976:375), the text reads *j-rotastak-n* and is translated as "in ... the capital (of the province of Ayrarat)."

Marr 1896:194 interpreted the word with reference to the Georgian passage. He compared Old Georgian *ro-j* with New Persian *rūd* "river." That it is in fact more probably from a northwest Iranian dialect is suggested, e.g., by Sulaimaniya Kurdish *îrō* "river" (Cabolov 1976:19), Talyš *rū* "river, stream" (Pirejko 1976:194). These, rather than New Persian *rūd*, cf. Middle Persian *rōt*, Old Persian *rautah-* n. and late Av. *θraotah-* n. "river" (in compounds, see Kellens 1974:81ff.), are more likely to be its source. Arm. *airu* "canal" is now understood as inherited (see Schmitt 1981:50, 72, but also Menges 1970:320f.). Even if we accept the inherited character of Arm. *airu* < **sruti-*, the prothetic vowel may be younger and a Pre-Armenian **iruy* cannot be fully excluded as the source of Old Georgian *ro-j/ru-j* "canal, stream" (see Ačaryan 1979). But to me the origin remains unclear, especially if we include the close Iranian words for "intestines," cf. e.g., Kurmanji Kurdish *rūvî* "kišečnyj kanal" (Farizov 1957:252) or Old Georgian *ruvi* "hollow for intestines" (Abulaze 1973:349), which has in certain Georgian dialects the same meaning as *ru*, but designates in other dialects a waterless channel in the mountains, where it is possible to descend (cf. Flonți 1975:64). Puzzling is also Arm. *ali-k'* "intestines" in the light of Osm. Turkish *ar(y)q* "trench for irrigation," etc. (for which see Menges, 1970:320f.)

In my view, Marr (1896:194) mistakenly connected the second element *stagi* with Persian *setāk*, which means "young shoot of a plant," and he wrongly interpreted it in this particular case as "vine." Accordingly he translated Armenian *rotastak* as "vinogradnik, orošaemyj kanalami" (1896:195), i.e., "vine(yard) to be irrigated by canals."

With this analysis Marr was close to a correct solution. However, he was also led astray, insofar as he explained the Armenian word on the basis of the Georgian passage and not vice versa. Now MP *lwst'k* [rōstak], later writted also *lwst'k* [rōstāk] "district, province" (Nyberg 1974:171) provides the *tertium comparationis* which leads to a simple solution. MacKenzie 1971:72 gives a more recent phonemic interpretation of *lw(t)st'k* as [rōstāg] "river-bed, district, province." Manichaean Parthian *rwdest'g* [rōdestāg] is rendered by Boyce 1977:79 as "province." I cannot follow Eilers' (1956:188 and 1977:295) etymological explanation as *rōta'stāk* "Flussort." In my opinion it has to be derived from an Old Iranian form **θrautas-tāk-* corresponding semantically to German "Flusslauf." We find it in late Av. *θraotō.stāc-* "qui court dans le flot" for which Kellens 1974:252f., 282f. has shown that it occurs four

times and always with long vowel. I understand the semantic development as “Flusslauf” > “Flussbett” > “Flussbezirk.”

In the Middle Persian text *Frahang i Pahlavīg* the term *rōtastāk* is an administrative subdivision which is smaller than *šahr* “dominion, kingdom” and larger than *dēh* “village” and *xānak* “house.” In New Persian we find *rūstā(i)* with the meaning “any inhabited place” and then in the sense of “a market-town, a village.” It has been taken over into Arabic as *rustāq* “arable land” (see Nyberg 1974:171), and Eilers 1977:281, 285, 295 has shown that later on it was often used as a placename.

As to the meaning of MP *rōtastāk*, note the following remarks by Marquart 1938:9 describing the old irrigation system in Iran:

“Es was in der Tat eine der wichtigsten Aufgaben der Satrapen oder Vizekönige des weiten Perserreiches, für die Instandhaltung dieses Bewässerungssystems, von welchem der Wohlstand und damit die Steuerkraft der ihrer Obhut anvertrauten Provinzen abhieng, Sorge zu tragen und die gleichmässige Verteilung des Wassers auf die *rōtstāk*-e zu überwachen.”

In conclusion, I would say that the Armenian text gives the correct form. The translation by Thomson, who follows Ačāryan 1979:146f., as “capital” may still be open to discussion. The Georgian text, however, has to be corrected. As we know, in Old Georgian the qualifying attributive tends to follow its noun with the so-called “long” case-ending of the genitive (with *a* added) and an anaphoric repetition of the case-ending of the noun qualified. Guided by these grammatical considerations, we see that the term *rudastag-ita* was thus through wrong separation of *ru* rendered *dastag-ita* (instr. case) *ru-ja* (gen. case) *-jta* (anaphoric instr. case). Instead of the reading

da dasxna kacni meruveni dastagita rujsajta

we therefore have to read:

*da dasxna kacni meruveni *rudastagita.*

We may translate as follows (taking the well-known ablative function of the Old Georgian instrumental case into consideration): “and he [Alexander] appointed people as canal overseers from the *Rudastagi*,” where *Rudastagi* is more probably the above-mentioned Iranian administrative term for territorial unit than a particular placename.

Deeters 1926:73ff. has shown that the Iranian voiceless stops were usually represented in Old Georgian as glottalized stops, but in later loans as aspirated stops: cf. e.g., Old Georgian *ṭalavari*, Arm. *taḷawar* “σκηνη” (Mt. 17:4, Luk. 9:33) and Manich. Parth. *tlw’r* [talawār] “hall,

tabernacle” but as a later loan Geo. *talari* “arbor, bower,” cf. NP *talār* “bed-chamber or saloon, built of wood and supported by four columns; a throne” (see Bielmeier 1985:35). Therefore we can take the placename *Rustavi*, a town not far from Tbilisi, as a loan from Early New Persian *rūstā* “village, market-place, etc.” containing the aspirated stop. This is derived from MP *rōtastāk*, from which also Old Georgian **rudastagi* in MK has come. We now see that Orbeliani’s translation of *sṭagi* as *rus satave* is nothing more than a folk etymology of *rusta(vi)* as *ru(j)s tavi*, lit. “head of the canal” i.e., “source of the canal.”

Finally we see that not only is the name of *Leonṭi Mroveli* derived from the possible Iranian loanword Old Geo. *ro(j)/ru(j)* “canal, stream” as *m-ro-vel-i* “the one belonging to the canal/stream,” but also that the name of the greatest Georgian poet *Šota Rustaveli*, and the name of the well-known Arabic geographer of the 9/10th century, *Ibn Rūstā*, are connected with the Early New Persian from *rūstā*, borrowed into Georgian with aspirated stop. Thus it is clear that *Rustaveli* is etymologically the primary form as against the also occurring form *Rustveli*.

Institut für Sprachwissenschaft der Universität Bern

References

- Abulaze, I. 1973. *Zveli kartuli enis leksikoni*. Tbilisi: Mecniereba.
- Ačariyan, Hr. 1979. *Hayeren Armatakan Baġaran, 4. hator*. Erevan: Erev. Ham. Hrat. Reprint of the Erevan edition, 1926-1935.
- Andronikašvili, M. 1966. *Narġevebi iranul-kartul enobrivi urtiertobidan, I*. Tbilisi: Tbilisis universiṭetiṣ gamomcemloba.
- Bailey, H. W. 1945. “Asica.” *TPhS* 1945:1-38.
- _____. 1985. *Indo-Scythian Studies being Khotanese Texts*. Vol. VII. London: Cambridge University Press.
- Bartholomae, Chr. 1904. *Altiranisches Wörterbuch*. Strassburg: Trübner. Reprint. Berlin, New York: de Gruyter, 1979.
- Benveniste, E. 1929. *Essai de grammaire sogdienne*. Paris: P. Geuthner.
- _____. 1945. “Deux noms divins dans l’Avesta [Srauša, Būšyaustā],” *RHR* 130, 1945: 13-16.
- Bielmeier, R. 1985. “Zu iranischen Lehnwörtern im Georgischen und Armenischen.” In *Sprachwissenschaftliche Forschungen. Festschrift für J. Knoblock*. IBK 23, eds. M. Ölberg, G. Schmidt, pp. 33-42. Innsbruck: Inst. für Sprachwiss. d. Univ.
- _____. 1990. “Sprachkontakt in der ‘Bekehrung Kartlis’,” In *Ägypten und Altes Testament, Lingua Restituta Orientalis, Bd. 20, Festgabe für J. Aßfalg*, pp. 30-44. Wiesbaden: Harrassowitz.

- Boyce, M. 1977. *A Word-List of Manichaean Middle Persian and Parthian*. Leiden, Téhéran, Liège 1977: E. J. Brill (=AcIr 9a).
- Cabолоv, R. L. 1976. *Očerki istoričeskoj fonetiki kurdskogo jazyka*. Moskva: Nauka.
- Deeters, G. 1926, 1927. "Armenisch und Südkaukasisch." *Caucasica* 3:37-82, 4:1-64.
- Eilers, W. 1956. "Der Name Demawend." *AcOr* 22:1954, 267-374, 24:1956, 183-224, 37:1969, 416-448. Reprint. Hildesheim, Zürich and New York: G. Olms, 1988.
- _____. 1977. "Prinzipien toponymischer Übertragung." *Onoma* 21:277-317.
- Farizov, I. O. 1957. *Russko-kurdskij slovar'*. Moskva: GIINS.
- Frye, R. 1969. "Continuing Iranian Influence on Armenian." In *Yad-Name-ye Irani-ye Minorsky*, ed. I. Afshar. Tehran = R. N. Frye, *Opera Minora*. Vol. I. The Pahlavi Codices and Iranian Researches 39, pp. 150-160. Shiraz: Asia Institute of Pahlavi University. 1976.
- Gershevitch, I. 1961. *A Grammar of Manichaean Sogdian*. Oxford: Blackwell.
- Gippert, J. 1984. "Zur historischen Onomastik des Georgischen." *Georgica* 7:37-42.
- Глonti, А. 1975. *Kartuli kilo-tkmata siqvis kona II*. Tbilisi: Ganatleba.
- Henning, W. B. 1937. "A list of Middle Persian and Parthian words." *BSOS* 9:79-92.
- _____. 1958. "Mitteliranisch." *Handbuch d. Orientalistik*. 1.Abtlg. 4.Bd. Iranistik. 1.Abschn. Linguistik, pp. 20-129. Leiden and Köln: Brill.
- Justi, F. 1895. *Iranisches Namenbuch*. Marburg: Elwert. Reprint. Hildesheim: G. Olms, 1963.
- KCx = *Kartlis Cxovreba*. S. Qauxčšvili (ed.). Vol. I. Tbilisi: Saxelgamo, 1955.
- Kellens, J. 1974. *Les noms-racines de l'Avesta*. Wiesbaden: Reichert.
- Kreyenbroek, G. 1985. *Sraoša in the Zoroastrian Tradition*. Leiden: Brill.
- Lommel, H. 1926. Review of Vasmer 1923. *Archiv f. slav. Philologie* 40:150-155.
- MacKenzie, D. N. 1971. *A Concise Pahlavi Dictionary*. London, New York and Toronto: Oxford University Press.
- Markwart, J. 1938. *Wehrot und Arang*, ed. H. H. Schaeder. Leiden: Brill.
- Marr, N. 1986. "Mnimoe geografičeskoe nazvanie *ṣrotastak* v Istorii Agaθangela." *Zapiski Vost. otd. Imp. russk. arxeolog. obšč.* 9:191-197.
- _____. 1902. "Bogi jazyčeskoj Gruzii po drevne-gruzinskim istočnikam." *Zapiski Vost. otd. Imp. russk. arxeolog. obšč.* 14:1-29.
- Mayrhofer, M. 1973. *Onomastica Persepolitana*. Wien: Verlag d. Österr. Akad. d. Wiss.
- _____. 1979. *Iranisches Personennamenbuch. Bd. I. Die altiranischen Namen*. Wien: Verlag d. Österr. Akad. d. Wiss.
- Menges, K. H. 1970. "Etymologika." In *W. B. Henning Memorial Volume*, eds. M. Boyce and I. Gershevitch, pp. 307-325. London: Lund Humphries.
- Miller, V. F. 1886. "Èpigrafičeskie sledy iranstva na juže Rossii." *ŽMNPb*, Oktjabr' 1886:232-283.
- _____. 1913. "K iranskomu èlementu v pripontijskix grečesksix nadpisjax." *IAK* 47:80-95.
- MK = *Mokceyav Kartlisay*, eds. B. Gigineišvili and El. Giunašvili, pp. 320-354. *Šatberdis krebuli X saukunisa*. Tbilisi: Mecniereba.

- Nyberg, H. S. 1974. *A Manual of Pahlavi. Part II: Glossary*. Wiesbaden: Harrassowitz.
- Orbeliani, Sulxan-Saba. 1966. *Txzulebani*. Vol. IV, 2 ed. I. Abulaze. Tbilisi: Sabčota Sakartvelo.
- Pätsch, G. 1975. "Die Bekehrung Georgiens. Mokcevey Kartlisay (Verfasser unbekannt)." *BK* 33:288-337.
- Pirejko, L. A. 1976. *Talyšsko-russkij slovar'*. Moskva: Russkij jazyk.
- Rastorgueva, V. S. 1966. *Srednepersidskij jazyk*. Moskva: Nauka.
- Schmitt, R. 1981. *Grammatik des Klassisch-Armenischen*. Innsbruck: Inst. f. Sprachwiss. d. Univ.
- _____. 1983. "Iranisches Lehnwort im Armenischen." *REArm* 17:73-112.
- Skalmowski, W. 1967. "Das Nomen im Parthischen." *Bull. Soc. Pol. de ling.*, fasc. 25:75-89.
- Steingass, F. 1892. *A Comprehensive Persian-English Dictionary*. London: Kegan Paul.
- Szemerényi, O. 1970. "Iranica III." In *W. B. Henning Memorial Volume*, eds. M. Boyce and I. Gershevitch, pp. 417-426. London: Lund Humphries.
- Takaišvili, E. 1900. "Istočniki gruzinskix letopisej. Tri xronika. 1. Obraščenie Gruzii (v xristianstvo)." *Sb. mat. dlja op. mest. i pl. Kavkaza* 28/1:1-116.
- Thomson, R. W. 1976. *Agathangelos. History of the Armenians*. Albany, N.Y.: State Univ. of New York Press.
- Vasmer, M. 1923. *Untersuchungen über die ältesten Wohnsitze der Slaven I. Die Iranier in Südrußland*. Leipzig: Veröff. d. balt. u. slav. Inst. d. Univ.
- Weber, D. 1986. "ΑΣΦΟΡΟΥΤΟΣ und Verwandtes." *Archäologische Mitteilungen aus Iran* 19:171-175.
- Zgusta, L. 1955. *Die Personennamen griechischer Städte der nördlichen Schwarzmeerküste*. Praha: Nakladatelství Českosl. Akad. věd.

Vowel Systems of Caucasian Languages⁰

J. C. Catford

Caucasian languages are widely supposed by general linguists to be characterized by possessing unusually large systems of consonants, balanced by unusually small systems of vowels. This characterization is only partly true, however, and is particularly applicable just to the Abkhazo-Adygan (Northwest Caucasian) languages. These languages, indeed, have vowel systems of only two or three items, and include large inventories of consonants — eighty in Ubykh, sixty-seven in the Bzyb dialect of Abkhaz, sixty-four in the Bzhedug dialect of Adyghe, though only forty-seven in literary Kabardian. Elsewhere in the Caucasus consonant systems are generally more modest, and vowel systems somewhat richer.

Though there is considerable variation in Caucasian vowel systems, there is some degree of zoning and before going on to more detailed discussion of some particular vowel systems, I would like to take an overall look at the geographical distribution of some characteristic features of vowel systems. Figure 1 is a kind of schematic map of the Caucasus, or better, a diagram showing the distribution of Caucasian languages and language groups in a spatial relation to each other that roughly matches their relative geographical locations. Figures 2 through 5 are plots of a few features of Caucasian vowel systems displayed on this same chart. The data used for making these charts is taken largely from Kibrik and Kodzasov (1978).

Figure 2 shows the distribution of three types of vowel system. All five NW Caucasian languages have unusual “linear” or, as I prefer to call them, “vertical” vowel systems in which phonologically pertinent oppositions occur only in the vertical dimension, that is, in the degree of aperture, all variations in the back-front dimension being purely allophonic. In Ubykh, Abaza and Abkhaz the system is of only two terms, usually represented as /ə/-/a/, while Adyghe and Kabardian have a three-term system, /ə/-/e/-/a/. It used once to be believed that these vertical vowel systems were unique to the Caucasus, but it is now known that similar systems occur in some languages of New Guinea (Pike 1964).

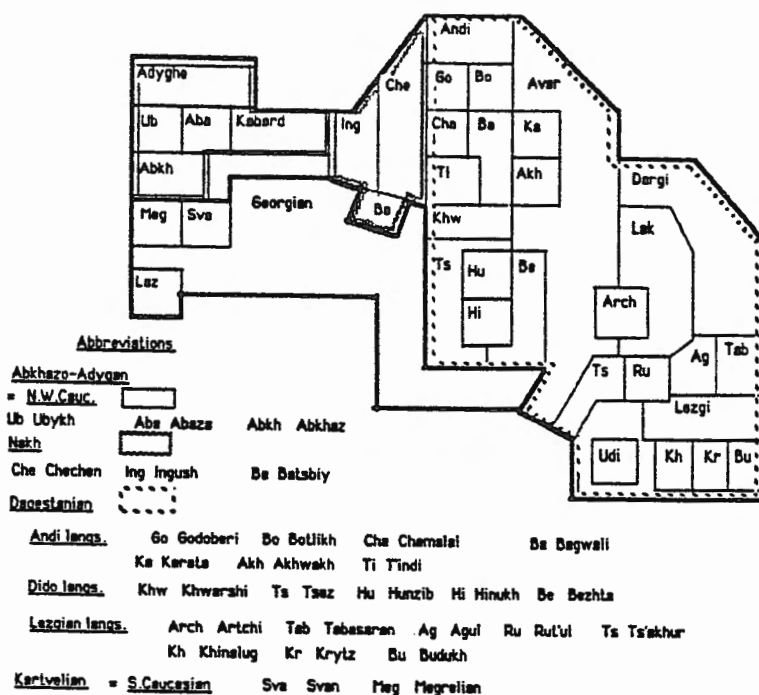


Fig.1 Caucasian Languages

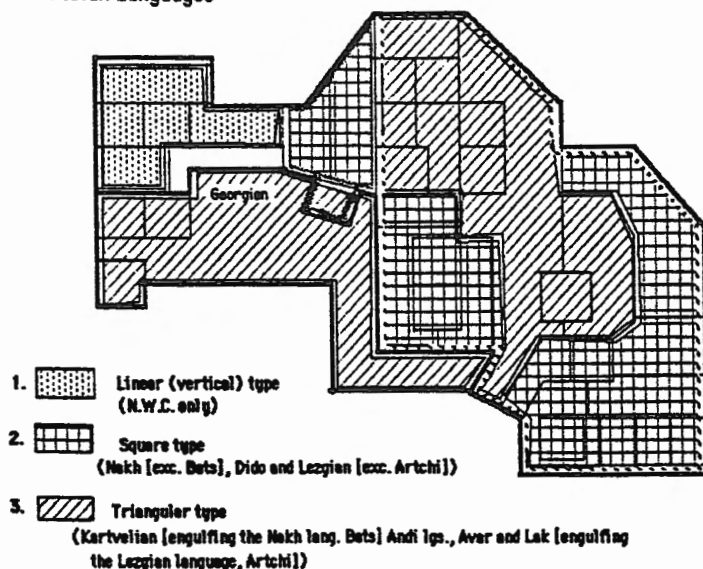


Fig. 2 Types of Vowel System

It is interesting to note that though unidimensional vertical vowel systems are attested, so far as I know no language has a unidimensional

“horizontal” system, in which there are phonological oppositions only between front and back vowels, any variations in the vertical dimension being allophonic.

The vowel systems of the other Caucasian languages are more orthodox, and are classified conventionally by Kibrik and Kodzasov as either **triangular** (in principle, with one maximally open vowel forming the triangle’s apex, contrasting with somewhat closer front and back vowels), or **square** (with no single open vowel as apex, but rather with parallel series of front and back vowels).

In some cases my own observations do not agree with those of Kibrik and Kodzasov — thus my recordings of Chechen seem to show a vowel system which might be better classified, as of the “triangular” type with three (unrounded) front vowels (two with long counterparts), three back vowels (two with long counterparts) and a fairly open, short/long vowel, thus front i/i: e/e: ε, back u/u: o/o: ɔ, and open, somewhat centralized a/a:.

In general, it is sometimes difficult to harmonize the conventional phonological description of vowel systems as “triangular” or “square” with the actual phonetic characteristics of the vowels concerned. We will look at the example of Avar in a moment.

Figure 2 shows that to some extent the typological character of vowel systems matches the (presumably) genetic subgrouping of Caucasian languages. Thus, vertical vowel systems are entirely confined to the five Abkhazo-Adygan (NWC) languages. The Kartvelian languages all have triangular vowel systems, a characteristic shared by Batsbiy, a Nakh language spoken entirely inside Kartvelian territory. In Dagestan, Avar, Lak, and the Andi languages all have triangular systems, while the Dido and Lezgian languages have square systems, with the exception of the somewhat isolated Lezgian language, Artchi, which, lying between Avar and Lak, has a triangular system.

Figure 3 shows the distribution of Caucasian languages which make phonological use of the opposition of short vs. long vowels. Among the NWC languages long vowels apparently occur in Ubykh and marginally in Abkhaz. All three Nakh languages have a clear-cut length opposition. Among the Dagestani languages, Avar and the northern Andi languages have no length opposition, but the southern Andi languages go along with the Dido languages, which, according to Kibrik and Kodzasov all have a length opposition with the exception of Khwarshi. However, according to Bokarev (1959) only Hunzib and Bezhta have fairly full sets of short/long vowels.

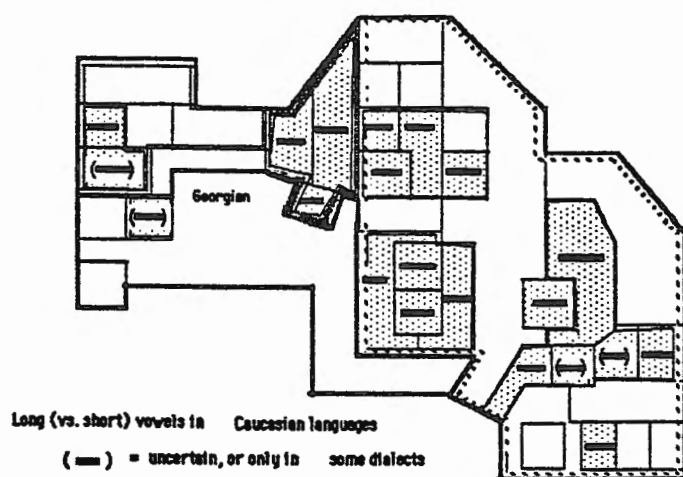


Fig. 3 Vowel length opposition

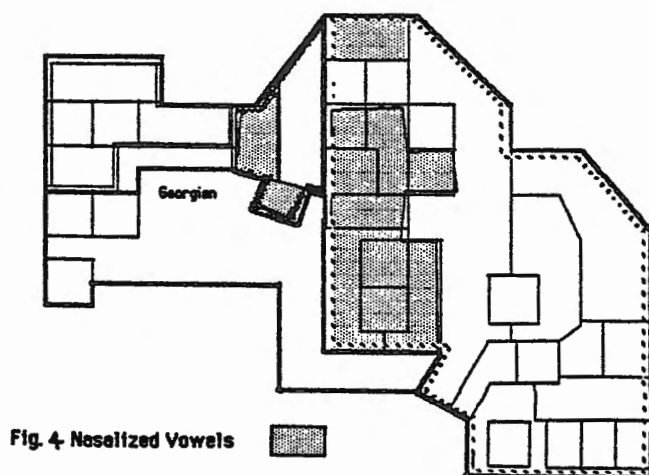


Figure 4 indicates that phonological nasalization of vowels is virtually confined to the Andi and Dido languages. These are languages with a simple syllable structure, with virtually no final consonants, a condition that is conducive to the development of nasalized vowels from sequences of vowel + nasal consonant.

Figure 5 shows that pharyngalization as a phonological feature of vowels is confined to Dagestan, though pharyngalized consonants are found in NWC, specifically in Ubykh and the Bzyb dialect of Abkhaz. In

some cases there is doubt whether pharyngalization is a property of vowels or of consonants. Thus Ibragimov (1978) claims that in the principal (Mukhad) dialect of Rutul, pharyngalization of vowels occurs only in the environment of pharyngalized consonants, though he had earlier (1968) described the pharyngalization of vowels as being independent of environment in Tsakhur. With respect to Tsakhur it seems to be the case that pharyngalized vowels form an independent series, as they also do in Udi. More on these vowels later.

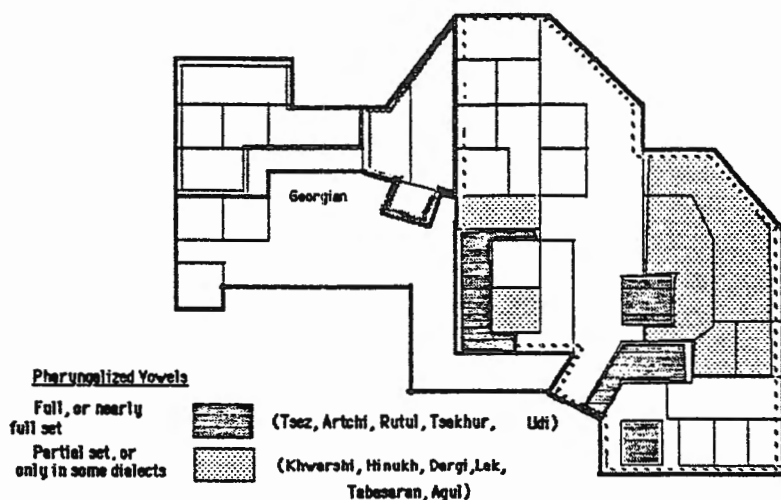


Fig. 5 Pharyngalized Vowels in Caucasian languages.

I would now like to look more closely at some particular examples of Caucasian vowel systems, beginning with the five Abkhazo-Adygan (NWC) languages. As is well known, the vowel systems of these languages have been the subject of much discussion. The question is: are any, or all, of these languages monovocalic? Jakovlev first broached the question in his 1923 detailed description of the phonetics and phonology of Kabardian, concluding that "...in the Kabardian language ... we can reliably establish only two short vowels, the phonemes *ǎ* and *ě*...." He goes on to point out that, noting the fugitive nature of *ǎ*, "...we must raise the question whether in Kabardian, at least in a certain period of its development, we have an example of a language with one vowel phoneme."

The starting point of the monovocalic hypothesis is the observation that, quite clearly, Kabardian vowels contrast in only one dimension. That is to say, that, though in spoken Kabardian one can certainly hear

vowel sounds occupying practically all of the total vowel space, it is very easy to demonstrate, by simple morphological commutation tests, that in fact these vowels all fall into three contrasting groups — a high group, represented by vowels of the types *i ɨ i# u u*, a mid group represented by *e ĕ ə ɔ̃ ʌ o*, etc. and a low type represented by *ä a*, etc., commonly represented as /ə/ /e/ /a/.

Nevertheless, Kabardian data certainly tend to suggest a monovocalic analysis, and, in fact, about 1948, before having seen Jakovlev's work, I had tentatively reached this conclusion, but could not substantiate it for lack of data. In 1960 Kuipers' book *Phoneme and Morpheme in Kabardian* carried the discussion a step further, purporting to show not only that Kabardian was, indeed, monovocalic, but that since its single vowel could be regarded as a "feature of openness" parallel to features of labialization and palatalization occurring on consonants, Kabardian was, in fact vowel-less.

This extreme position was criticised by Halle (1970), and is not generally accepted. I have elsewhere (Catford 1984) explained in detail why I no longer accept the monovocalic hypothesis, and regard Kabardian as having a vertical three-vowel system, which, following Trubetzkoy (1925), we may represent as /ə-e-a/. Nevertheless I think it is worth while recapitulating some of the discussion here.

The monovocalic hypothesis depends upon the elimination of both the high vowel phoneme /ə/ and the low vowel phoneme /a/, leaving only the mid vowel phoneme /e/. In Kuipers' notation, these are transcribed /ə/ /a/ and /ā/, since he takes the low vowel phoneme to be a long vowel representing an underlying /a/ (the mid vowel in his notation) plus a covert /h/. This double elimination, as I point out in the 1984 article, can only be done by taking some liberties with the data.

In the first place one can maintain the view that /ə/ is not a phoneme, but a predictable interconsonantal, or postconsonantal support vowel only by analysing initial consonant clusters in what seems to be a highly artificial way, and by disregarding the fact that there is at least one other environment in which the interconsonantal opposition /ə/-/zero/ is significant.

Consonant clusters such as /pʰ/ /bʰ/ /p'ʂ'/ must be distinguished from the corresponding sequences /pəʰ/ /bəʰ/ /p'əʂ'/ in such contrasting pairs as:

/pʰaʂ/	'he looked'	/pəʰaʂ/	'it hung'
/bʰləm/	'seven' (oblique case)	/bəʰləm/	'cattle'
/p'ʂ'en/	'to weed out'	/p'əʂ'en/	'to freeze'

Kuipers eliminates such cases of phonological contrasts between /zero/ and /ə/ by attributing the differences between such pairs to the initial

consonant, setting up, in effect, a special set of “close transition consonants” contrasting with those having a brief ə-type open transition to the second consonant.

The second unjustifiable elimination of the /zero/-/ə/ opposition concerns the alternation of forms with and without /ə/ as the subject pronominal prefix in the past tense of verbs. Here, intransitive verbs have the prefixes *sə- wə- də- fə-*, whereas transitives have the vowelless prefixes *s/z- p/b- t/d- f/v-*, as in these examples:

/sə-txaŝ/	‘I wrote’	(intrans)	/s-txaŝ/	‘I wrote’	(trans)
/fə-txaŝ/	‘you wrote’	(intrans)	/f-txaŝ/	‘you wrote’	(trans)
/sə-daŝ/	‘I sewed’	(intrans)	/z-daŝ/	‘I sewed’	(trans)

The elimination of the open vowel (/a/ in my notation, /ā/ in Kuipers’) rests on the assumption that this open vowel contrasts with the mid vowel /e/ (Kuipers’ /a/) in **quantity** rather than quality — and can thus be interpreted as /e/ (Kuipers’ /a/) plus length, or accompanied by a, usually covert consonant. The specific consonant posited for this function is /h/, and this is justified on two grounds.

First, the open vowel is the only vowel occurring initially in native Kabardian words, such as /ade/ ‘father’, /adəʔe/ ‘Adyghe’, and in this position it is preceded in the pronunciation of some Kabardian speakers by a murmured fi. Therefore, the argument goes, initial *a* can be analysed in all cases (even in the pronunciation of those speakers who do not pronounce the fi) as underlying /h/ plus /e/ (Kuipers’ /a/). Doubt is cast on this hypothesis by the fact that, though /a/ is the only initial vowel in Kabardian words, nevertheless, there is another circumstance in which literate Kabardians also pronounce initial /ə/ and /e/, namely in citing the letters ы, э, and а of the (Cyrillic) Kabardian alphabet. Here all three vowels occur, and all may be preceded by fi: thus fiə, fie, and fia are kept quite distinct, and so one cannot regard initial /a/, pronounced fia, as “really” /he/. It appears, in other words, that murmured fi is simply a Kabardian way of producing a “soft attack” on a vowel (in contrast to the “hard attack” of glottal stop ʔ, which is phonemic in Kabardian).

Secondly, Kabardian has a post-accentual plural suffix realised as /xe/ in literary Kabardian and in about half the Kabardian dialects as well as all dialects of Adyghe. In some Kabardian dialects this plural morpheme is realised as /he/, quite distinct from /xe/, which also occurs in these dialects. Kuipers equates the phonological form of this suffix, which is always pronounced quite distinctly as *xe* (*he* in some dialects) with initial /a/. He then identifies morphemes with the pre-accentual plural pronominal morpheme /a/, analysed as /he/ or /eh/ (that is, Kuipers’ /ha/ or /ah/). It is particularly difficult to accept this analysis for literary

Kabardian, where the plural suffix is realized as /xe/ and the syllable /xe/ can also contrast initially with /a/ (/ā/), as in /ade/ 'father', /xede/ 'choose'.

The stressed /a/ (ā) marking the past tense of verbs is also analysed as underlyingly /eh/ (Kuipers' /ah/). Thus all 'long' vowels can be regarded as having the same underlying structure, i.e., V + glide (-w, -y, or -h). However, in the case of final *i*, *u*, *e*, *o* it can easily be demonstrated by simple morphological tests that these are underlyingly /əy/ /əw/ /ey/ /ew/, whereas, for /a/ (/ā/) no such tests are available and the analysis is thus unverifiable. It is true that the fact that this final /a/ (/ā/) is stressed suggests that a consonant may have followed it in an earlier stage of the language, but that scarcely justifies the analysis /ah/ for the contemporary language, and in any case the formerly present consonant was probably not *h* but a voiced uvular fricative ʁ, as in contemporary Adyghe, and, according to Kumaxova and Kumaxov (1979:82) still occasionally encountered in Kabardian oral folk poetry.

Further doubt is cast on the association of 'long' /ā/ with the other 'long' vowels by textual frequency statistics. Balkarov (1970) tabulates the frequency of occurrence of Kabardian phonemes in 1000 words of text. The relevant frequencies are:

ə	617
e	876
a (=eh?)	413
ē (=ey)	37
ō (=ew)	16

These figures show that the frequency of /a/ (/ā/) is an order of magnitude greater than that of the 'long' vowels, and that /a/ is thus more closely associated with the 'short' vowels, /ə/ and /e/.

Finally, the assumption that /a/ is underlyingly /eh/ or /he/ rests in part, as I pointed out above, on the assumption that the open vowel /a/ contrasts with the mid vowel in **quantity** rather than **quality**. Average duration of Kabardian vowels, for one speaker, are as follows: /ə/ 92 ms., /e/ 117 ms., /a/ 143 ms. That is to say, the average duration of /e/ is 82% of that of /a/. This would be an abnormally high percentage ratio for a short/long vowel pair, where a ratio around 50% is normal. It is more likely to be the normal duration difference between a closer and an opener vowel.

Moreover, from my data it is clear that there is practically always a well marked difference in quality between tokens of /e/ and those of /a/. Figure 6 shows a Cardinal Vowel + Formant chart, that is a Cardinal Vowel diagram with values of the first two vowel formants incorporated – obviously different formant values apply to rounded and to unrounded

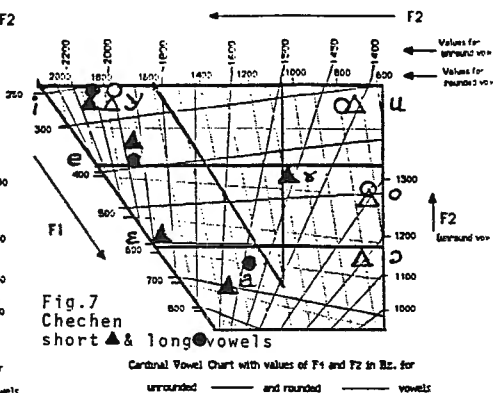
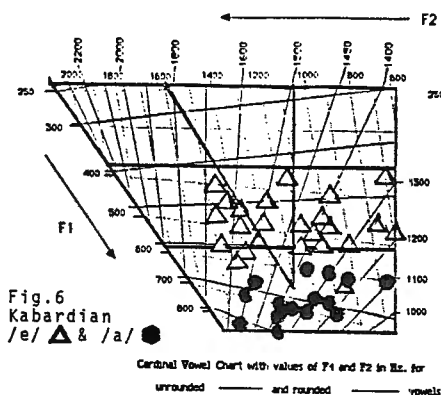
vowels, which complicates the figure somewhat. Using formant frequencies extracted from recordings of one speaker, some fifteen tokens of Kabardian /a/ and twenty-four of /e/ are plotted. It can be seen that with one exception the formant frequencies (and hence the acoustic/auditory quality) of /a/ are quite distinct from those of /e/.

It seems clear to me that the Kabardian vowel system is best described as a vertical system of three vowels /ə/ /e/ /a/. It is even more clear, on the basis of numerous contrasting sets of items that the vowel system of Adyghe is of the same type, though the other three NWC languages have only two vowels. As M. A. Kumaxov showed in his 1973 article, the vowel systems of NWC languages are:

Abkhaz, Abaza, Ubykh
Kabardian, Adyghe

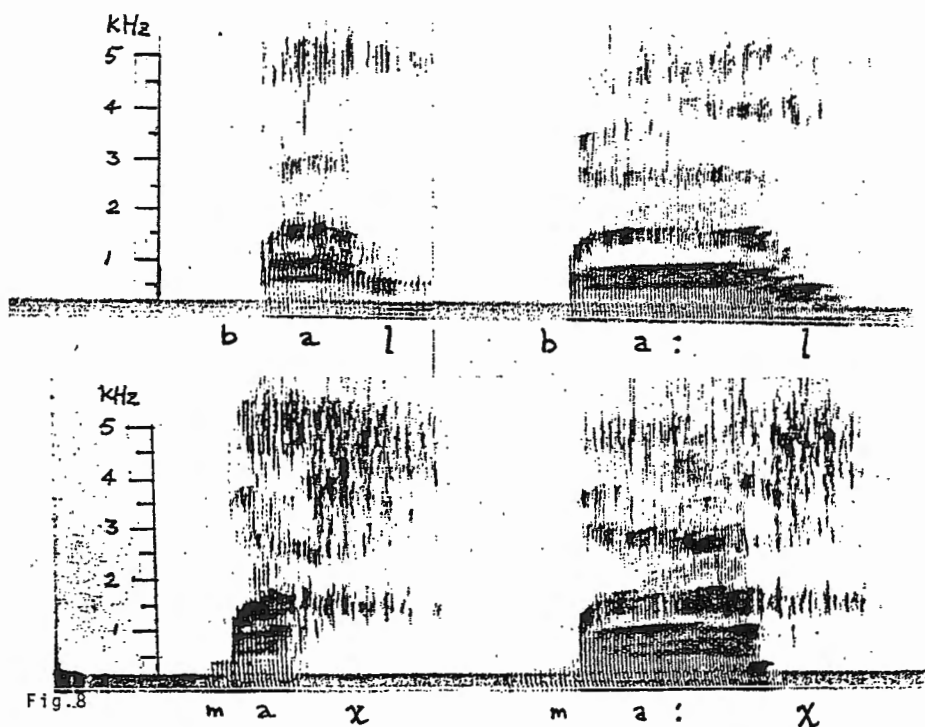
ə-a
ə-e-a.

Having shown that there is probably no phonological length in the Kabardian vowel system, I turn now to the Nakh languages, and specifically to Chechen, where quantity certainly plays a role in the vowel systems of all three languages, Chechen, Ingush, and Batsbiy. Chechen presents the greatest contrast to the NWC languages, since it has been described as having no fewer than thirty vocalic nuclei, fifteen monophthongs and fifteen diphthongs.



The fifteen monophthongs resolve themselves into nine short vowels and six long vowels. This, at least, is the vowel inventory of Chechen as described in Deśeriev (1960) and as recorded by me in Moscow, from an informant claiming to speak the Plains dialect upon which the literary language is based. The duration ratios of short to long vowels for the

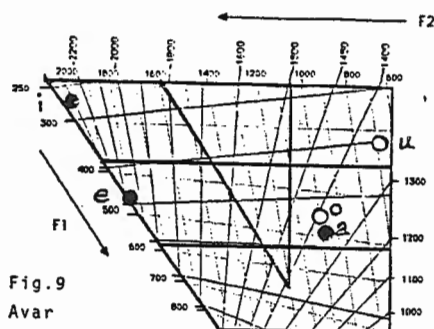
Chechen vowel pairs in my data are /i/-i:/ 50%, /e/-e:/ 46%, /a/-a:/ 49%, /o/-o:/ 52%, /u/-u:/ 42%, /y/-y:/ 48%, a much more 'normal' ratio for short vs. long vowels, where the duration difference is truly significant, than the 82% for Kabardian /e/-/a/ reported above. The qualities of the Chechen short vs. long vowels are quite remarkably similar, as is shown in Figure 7, where the Chechen vowel pairs are plotted on a Cardinal Vowel + Formants chart, short vowels being shown by triangles, long vowels by circles — open triangles and circles indicating rounded vowels.



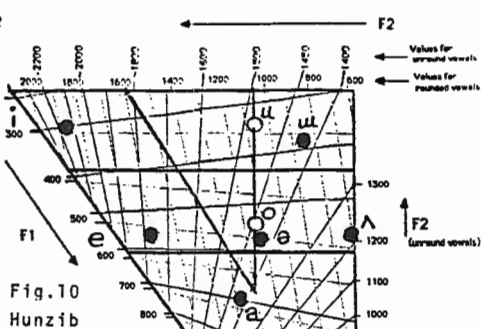
The validity of this analysis of Chechen vowels has been called into question — in particular, it has been suggested that the short vs. long opposition for the open vowels /a/-/a:/ is allophonic. At one level, however, it is clearly significant for my informant. Figure 8 shows spectrograms of the word pairs /bal/ 'cherry tree' and /ba:l/ 'burden', and /max/ 'price' /ma:x/ 'needle'. It is clear from these spectrograms that the pronunciation of the items in each pair is essentially identical, except for the length of the vowel. However, the orthographic forms of these words

are балл and бала, and мах and маха respectively, and this suggests possible differences in the underlying forms.

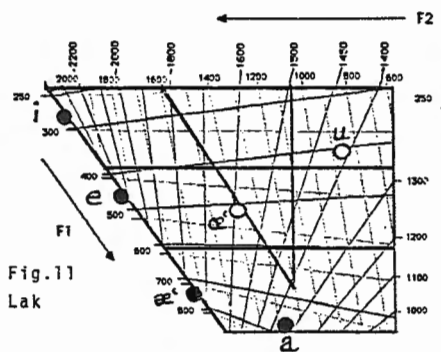
Figures 9 through 13 are examples of some Dagestanian vowel systems, each showing the locations of the vowels on a Cardinal Vowel + Formants chart, with an indication, in the form of a triangle or a square, of the type of vowel system as classified by Kibrik and Kodzasov (1978). The formant values of the vowels, and hence their locations on the charts, are in fact means derived from varied numbers of recorded examples. On all charts, filled circles represent unrounded vowels, unfilled circles represent rounded vowels.



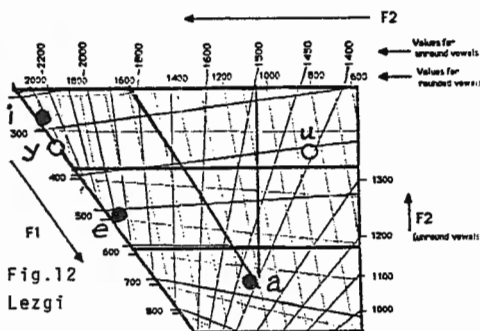
Cardinal Vowel Chart with values of F1 and F2 in Hz. for unrounded — and rounded — vowels



Cardinal Vowel Chart with values of F1 and F2 in Hz. for unrounded — and rounded — vowels



Cardinal Vowel Chart with values of F1 and F2 in Hz. for unrounded — and rounded — vowels



Cardinal Vowel Chart with values of F1 and F2 in Hz. for unrounded — and rounded — vowels

Figure 9 shows the vowels of literary Avar, a supposedly triangular system, but it is noteworthy that the "apex" of the triangle, the vowel /a/ has almost the same location on the vowel diagram as the vowel /o/,

which differs from it chiefly in terms of lip rounding. It is interesting to note that tracings of X-ray images of Avar /a/ and /o/, when superimposed, show almost identical tongue positions (Žaparije, 1967). Figure 10 shows the very unusual vowel system of the Dido language Hunzib. This language has a system of eight short vowels, of which five are fairly "orthodox" /i/ /e/ /a/ /o/ /u/, except that none are peripheral (a rather common feature of Caucasian vowel systems), the rounded vowels /o/ and /u/ being virtually central; the remaining three vowels are more unusual, being central or back unrounded vowels, /ə/, /ʊ/, and /ʌ/, represented in Cyrillic transcription as ə, ʊ, and а.

The Lak vowel system, Figure 11, consists of three plain vowels /i/ /a/ /u/ (realized as an under-rounded ʊ-like vowel), and three slightly pharyngalized vowels, /e'/ /æ'/ /œ'/, the last being highly centralized.

Lezgi, Figure 12, has a vowel system somewhat similar to that of Lak, except that in place of the Lak centralized /œ/ Lezgi has a fully front, rather close /y/.

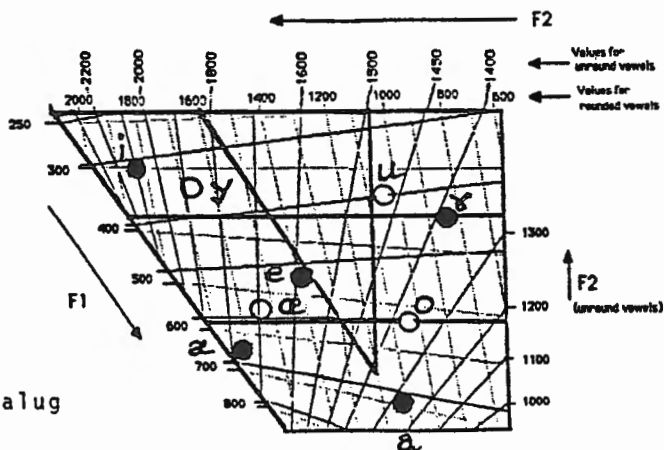


Fig.13 Khinalug

Our final Daghestanian example in this set is Khinalug, Figure 13. Khinalug has both two rounded rather front vowels, /y/ and /ø/, and an unrounded half-close back vowel /ɤ/. As the vowel chart shows, Khinalug exhibits rather strongly the Caucasian tendency to have nonperipheral vowels.

Figures 14 and 15 exemplify two Kartvelian languages. Figure 14 shows the simple triangular five-vowel system of Georgian, exhibiting, once again, the tendency for vowels to be non-peripheral. Figure 15 represents the vowel system of a speaker of the Lower Balsk dialect of

Svan. This system differs from that of Georgian by the inclusion of a low front vowel /æ/ and a half-close unrounded back vowel /ɤ/.

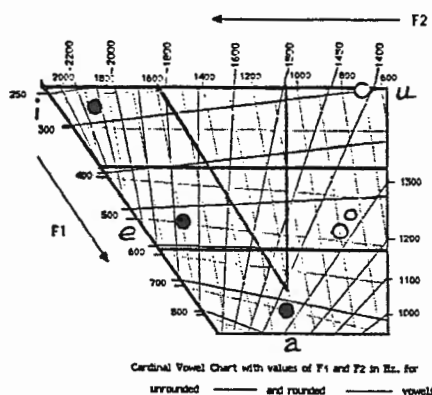


Fig. 14 Georgian

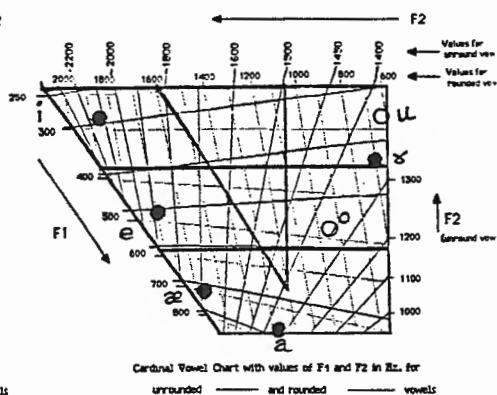


Fig. 15 Svan, Lower Balsk dialect

Some dialects of Svan have an opposition of vowel length, but not Lower Balsk. On the other hand this dialect exemplifies in a certain way one of the most striking features of Svan, namely the existence of the results of a process of umlaut. Soviet Caucasologists often describe certain vowels of Caucasian languages (vowels of the types æ y œ) as “umlauted,” even when they are not due to the particular diachronic process known as umlaut in western Indo-European languages. In Svan, however, as in some dialects of Chechen, a genuine umlaut process — stressed vowel modified by quality of unstressed vowel in following syllable (which may subsequently be lost) — has occurred. In some dialects this has apparently led to the presence of rounded front vowels (represented by *ü* and *ö* in the literature) as well as the low front vowel /æ/ (*ä*). This latter vowel occurs, for example, as the result of a-umlaut in /didæb/ ‘fame’, cf. Geo. /dideba/. However in Lower Balsk there is no y or œ. The i-umlauted reflexes of *u* and *o* are usually /wi/ and /we/, as in /pwir/ ‘cow’, pronounced approximately as p̥i̯r (the /w/ being palatalized by the following vowel, and at least partly devoiced by the strong aspiration of the /p/) cf. Geo /puri/, and /twep/ t̥w̥ɛpʰ ‘gun’, Georgian /topi/.

The vowel system of the Megrelian dialect of Zan is very similar to that of Lower Balsk, minus the low front vowel /æ/.

I return now to Dagestanian to describe the interesting type of pharyngalization of vowels in Udi and Tsakhur. Each of these languages has a set of what are called ‘pharyngalized’ vowels. I begin by listing the

average formant values of the first, second, and third formants of the vowels of Tsakhur the Vartashen dialect of Udi.

TsakhurUdi

	F1	F2	F3		F1	F2	F3
i	297	2480	3057	i	273	2233	2867
i ^ϕ	388	2181	2840	i ^ϕ	300	1950	2573
e	485	2250	2800	e	473	2150	2800
e ^ϕ	548	2063	2601	e ^ϕ	535	1650	2300
a	729	1461	2422	a	724	1000	2386
a ^ϕ	738	1653	2066	a ^ϕ	648	1430	1900
o	533	1044	2900	o	400	775	?
o ^ϕ	567	1220	1825	o ^ϕ	450	1150	1600
u	347	1000	3000	u	300	700	2400
u ^ϕ	500	1125	1850	u ^ϕ	350	1050	2050

Udi also has three vowels not subject to pharyngalization, with mean F1 and F2 as follows: /æ/ 675:1867, /y/ 290:1800, /ø/ 460:1760)

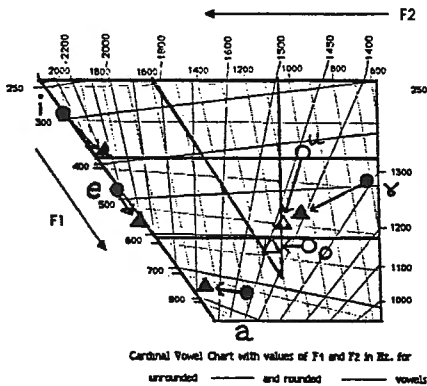


Fig.16 Tsakhur

Plain ● & pharyngalized ▲ vowels

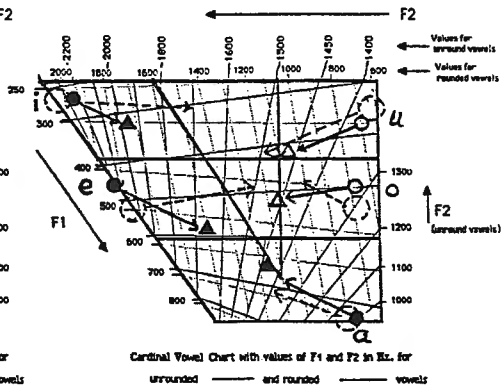
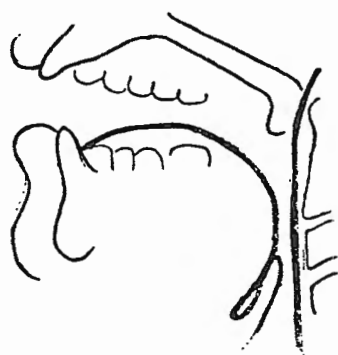


Fig.17 Udi, Vartashen dialect

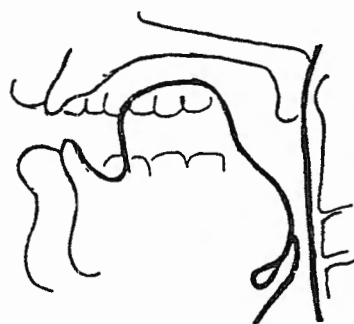
Spectrographically identified vowels ●, Perceptually (auditorily) identified vowels ○.

Figures 16 and 17 show these vowels of Tsakhur and Udi plotted on Cardinal Vowel + Formant charts, with the pharyngalized vowels

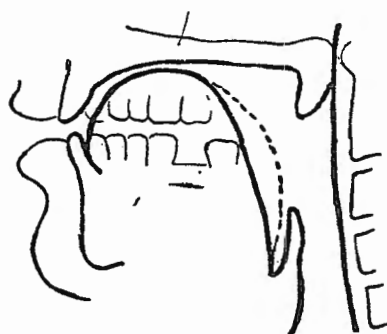
represented as triangles joined by arrows to the plain vowels. It is clear from these that the pharyngalized vowels acoustically resemble vowels that have been modified in specific ways. Usually the back vowels appear to have been centralized and the front vowels lowered and/or centralized. This, indeed, is the auditory effect produced by these vowels, as is shown by the large broken circles and broken arrows on the Udi chart. These represent the auditory assessments I made of the Udi vowels before I had acoustic data from recordings. Clearly, the auditorily perceived shifts correspond fairly well to the actual acoustic shifts.



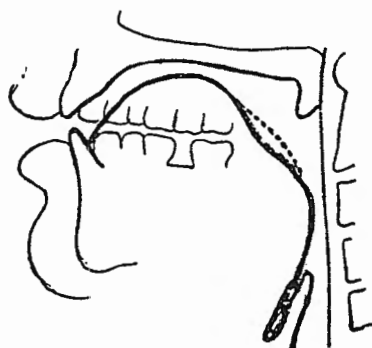
Udi /a/



.Udi pharyngalized /aʁ/



Tsakhur /i/



Tsakhur pharyngalized /iʁ/

Fig.18 Tracings from X-rays of tongue configurations for plain and pharyngalized vowels in Udi and Tsakhur (from Gaprindašvili, 1966)

In fact, the actual articulation of these vowels is known from the X-ray tracings presented in Gaprindašvili (1966), from which I take the examples in Figure 18. These examples show the tongue configurations of Udi /a/ and /a[̣]/ and of Tsakhur /i/ and /i[̣]/. It is quite clear from these that there is indeed pharyngalization — the root of the tongue, with the epiglottis, can be seen projecting backwards into the pharynx. But at the same time there is sulcalization, or hollowing, of the dorsal surface of the tongue approximately opposite the uvula.

Now this configuration is remarkably similar, as I have pointed out elsewhere (Catford 1983), to a tongue shape observed in a common variety of Midwestern American /r/ described, with X-ray data, first by Delattre and Freeman (1968) and more recently by Ladefoged (1979). It is also interesting to note that these pharyngalized vowels have a measurable acoustic feature in common with rhotacized vowels, namely the lowering of F3 as compared with the corresponding plain vowel. This is noticeable for all the Tsakhur and Udi vowel pairs with the exception of the Udi /o/, for which I was unable to get an accurate reading of F3. Because of this similarity between these pharyngalized vowels and rhotacization it is not surprising that Americans (so I was told by the Moscow phonetician, S. Kodzasov) think they hear an /r/ in the sound of these vowels.

Caucasian languages, as I have shown, have a considerable range of different types of vowel system, of which the most unusual and interesting ones are the vertical vowel systems of the Abkhazo-Adygian (NWC) languages, and the systems with pharyngalized vowels in Dagestan.

University of Michigan

Note

⁰The research upon which this paper is based was supported in part by National Science Foundation Grant No. BN 877-16909. Opinions expressed in this publication are those of the author and do not necessarily reflect the views of the NSF.

References

- Balkarov, B. X. 1970. *Fonetika adygskix jazykov*. Nalčik: Elbrus.
 Bokarev, E. A. 1959. *Cezskie (didoiskie) jazyki Dagestana*. Moscow: Akad. nauk SSSR.

- Catford, J. C. 1983. "Pharyngeal and Laryngeal Sounds in Caucasian Languages." *Vocal Fold Physiology*. Diane M. Bless and James H. Abbs (Eds.), pp. 344-350, San Diego: College Hill Press.
- _____. 1984. "Instrumental Data and Linguistic Phonetics." *Topics in Linguistic Phonetics in Honour of E. T. Uldall*. Jo-Ann Higgs and Robin Thelwall (Eds.), pp. 23-48. New University of Ulster: Occasional Papers in Linguistics and Languages Learning No. 9.
- Delattre, P. and D. C. Freeman. 1968. "A dialect study of American R's by X-ray motion picture." *Linguistics* 44:29-68.
- Dešeriev, Ju. D. 1960. *Sovremennij čečenskij literaturnyj jazyk. Čast' 1. Fonetika*. Groznyj: Čečeno-inguškij naučno-issledovatel'skij Institut istorii, jazyka i literatury.
- Ĵaparije, Z. N. 1967. "Salit'erat'uro xunjuri enis xmovnebi." *Met'qvelebis analizis, sintezisa da st'at'ist'ik'is sak'itxebi*. 112-21. Tbilisi, Mecniereba.
- Gaprindašvili, S. G. 1966. "Xmovanta paringalizaciis bunebisatvis c'axurulsa da udur enebši." *Met'qvelebis analizis da sintezis sak'itxebi* 28-80. Tbilisi, Mecniereba.
- Halle, M. 1970. "Is Kabardian a vowelless language?" *Found. Lang.* 1:11-44.
- Ibragimov, G. X. 1968. *Fonetika caxurskogo jazyka*. Maxačkala: Dagestan. filial, Akad. nauk SSSR.
- _____. 1978. *Rutul'skij jazyk*. Moscow: Nauka.
- Jakovlev, N. F. 1923. *Tablicy fonetiki kabardinskogo jazyka*. Moscow.
- Kibrik, A. E. and S. V. Kodzasov. 1978. "Fonetičeskie obščnosti." *Strukturnye obščnosti kavkazskix jazykov*. 86-118. Moscow: Nauka.
- Kuipers, A. H. 1960. *Phoneme and Morpheme in Kabardian*. The Hague, Mouton.
- Kumaxov, M. A. 1973. "Teorija monovokalizma i zapadnokavkazskie jazyki." *Voprosy jazykoznanija* No. 6:54-67.
- Kumaxova Z. J. and M. A. Kumaxov. 1979. *Funkcional'naja stilistika adygskix jazykov*. Moscow: Nauka.
- Ladefoged, P. 1979. "Articulatory Parameters." *UCLA Working Papers in Phonetics* 45:25-31.
- Pike, E. V. 1964. "The Phonology of New Guinea Languages." *American Anthropologist* 64 (4):121-32.
- Trubetzkoy, N. S. 1925. Review of Jakovlev *Tablicy fonetiki kabardinskogo jazyka*. *BSL* 25:277-81.

How to Describe the Sounds of the Northwest Caucasian Languages

John Colarusso

1. Introduction

1.1. The languages of the Northwest Caucasian family fall into three branches: Abkhaz-Abaza with Bzyb, Abzhwi, Sadz, Samurzakan and Ahskharawa dialects of Abkhaz and Abaza (with several sub-dialects); Ubykh (with two dialects, one very poorly attested); and Circassian, with two sub-branches: the Eastern with Kabardian, Besleney and Kubano-Zelenchuk, and the Western with Chemgwi, Yegerukay, Hakuchi, Bzhedukh, Shapsegh, Natukhay and Abadzhakh. As the name of the family indicates these languages were native to the northwestern region of the Caucasus mountain area, i.e., they ran along the eastern shore of the Black Sea, from the Sea of Azov in the north almost down to the Turkish border and from the coast inland, over the mountains roughly half way to the Caspian. In 1864, after the Tsarist conquest of the region (Baddeley 1908), many speakers left for the Ottoman Empire so that today there are large populations of the speakers scattered throughout the Middle East, with smaller pockets in the Balkans, and even a few more recent immigrants elsewhere in Europe and in North America.

1.2. The languages are justly famous for their extraordinary sound systems. Not only do they exhibit a phenomenal degree of consonantism, but this is coupled with what may be near minimal vowel systems. Kabardian has the fewest consonants and glides with forty-nine, while Ubykh has eighty-one and Abaza may be analyzed as having eighty-three (Colarusso 1975:213-17). No one who has done any substantial work on the languages has ever posited more than three vowels, usually written as either /ə, e, a/ or /ə, a, ʌ/. In Colarusso (1982) I have shown that the systems are actually /ə, a/ with the second vowel, which is distinguished only by being [+low], subject to lengthening (and concomitant lowering) when followed by /a/ or /h/ in the Circassian languages, /a/ in Ubykh, or underlying /ʕ/ in Abkhaz. Abaza shows only the two vowels, significantly lacking any of the phonological processes responsible for the third surface vowel found in the other languages. As odd as such a vocalic system may appear, it can be found elsewhere: in the Ndu family of New Guinea (Laycock 1965; Pike 1964), in Chinese (Colarusso 1975:378-79), in the Ethiopic Ge'ez (ibid., pp. 380-81), and perhaps in

Proto-Indo-European (Colarusso 1981:499-502; Lehmann 1955) and Proto-Kartvelian to the south of the Northwest Caucasian languages (Gamkrelidze 1966; Gamqrelize and Mačavariani 1965).

1.3. In the Northwest Caucasian languages this near-minimal vocalic system is complementary to the consonantal inventories. Phonetically there is a wide range of vowel colors. In these languages such colorings, however are not interpreted as inherent to the syllabic peak, but rather as being due to the effects of the syllabic margins, with syllable offset taking precedence over onset where possible (Colarusso 1983). The only feature that cannot be attributed to consonantal margins (apart from the syllabic peak itself (Colarusso 1982)) is that of [+low], whose opening gesture is antagonistic with the radical closure necessary for all but laryngeal and perhaps pharyngeal sounds. Thus the very paucity of the vowel inventories in these languages gives rise directly to extremely large consonantal ones.¹ One finds rounded, pharyngealized, pharyngealized and rounded, palatalized ([+high] and [+Advanced Tongue Root]), labialized and labiodentalized consonants. Furthermore, most of the languages make use of every possible point of articulation: labial (and labiodental), dental, alveolar, alveolopalatal, palatoalveolar, palatal, velar, uvular, pharyngeal, and laryngeal. While one can find languages that have more consonantal phonemes than do the Northwest Caucasian², nowhere else can one find anything like the combination of such extensive use of the vocal tract's articulatory points with secondary modifications of such diversity.

2. Purpose of this Study

2.1. Given the unique phonological and phonetic richness of these languages³ they present a unique opportunity to the phonetician to examine the acoustic potential of man within the confines of one language or even one speaker. The present study gives a brief summary of phonetic research soon to appear (Colarusso, in preparation) accompanied by extensive spectrographic material. Present space requirements have led to the use of graphs to represent salient features of points of articulation as they couple with various secondary modifications and source features. Such work is of interest to the phonetician because it permits a study of these consonantal properties free from effects of articulatory assimilation to neighboring vowels. Further, as only two speakers are presented (Mr. Tefvik Esenç for Ubykh and Mr. Rashid Dahabsu for Bzhedukh West Circassian), one can study the partitioning of acoustic space without the smearing due to statistical variation within a body of informants. The elimination of such smearing is important if one seeks a quantal theory of phonetic perception (Stevens 1972): a statistical distribution may enable one to determine the centers of critical acoustic

ranges, but will obscure the boundaries of these same zones: whereas the examination of one or two speakers with similar acoustics (the case with Esenç and Bahabsu, luckily enough) permits the delimitation of such boundaries while at the same time suggesting likely centers. The graphs represent not only ranges, but such possible centers as well.⁴

2.2. This study also tries to clear the air for the phonologist and specialist in Caucasian studies. Given the subtlety and complexity of these languages' sound systems it is not surprising that much debate and confusion have centered around the basic inventories of many of the languages, (cf., Leroy and Paris [1974] for a summary of the debate surrounding Ubykh). One can even find otherwise excellent works in which crucial phonological contrasts are totally ignored, (for example, Paris [1974a] where /š^{hy}/ ~ /šy/ and /χ^h/ ~ /χ/ are not observed), or misrepresented, (as in every work that mentions Abaza, except that of Bgažba [1964], which all fail to note the pharyngealization of the aspirated uvulars).

2.3. Thus I herein proceed from the lips back through the vocal tract to the larynx. Each point of articulation is named, along with the segments found thereat, noting restriction by languages. A phonological characterization is given in terms of distinctive features (Colarusso 1979; Perkell 1971; Klatt and Stevens 1969; Halle and Stevens 1969; Chomsky and Halle 1968:293-329). Wherever possible mid-sagittal x-ray tracings are given. These are for Ubykh (Leroy and Paris 1974) and Bzyb Abkhaz and Abaza (Bgažba 1964). Points of phonetic or phonological worth are briefly discussed. For clarity modified series are treated separately from homorganic series that are plain or bear some other modification, e.g., /q^h, q, q', x, x', γ/ is presented separately from /q^{hw}, q^w, q'^w, x^w, x'^w, γ^w/ even though both are uvular series.

3. Labials /p^h, p, b, p', f, v, f', m, w, w'/

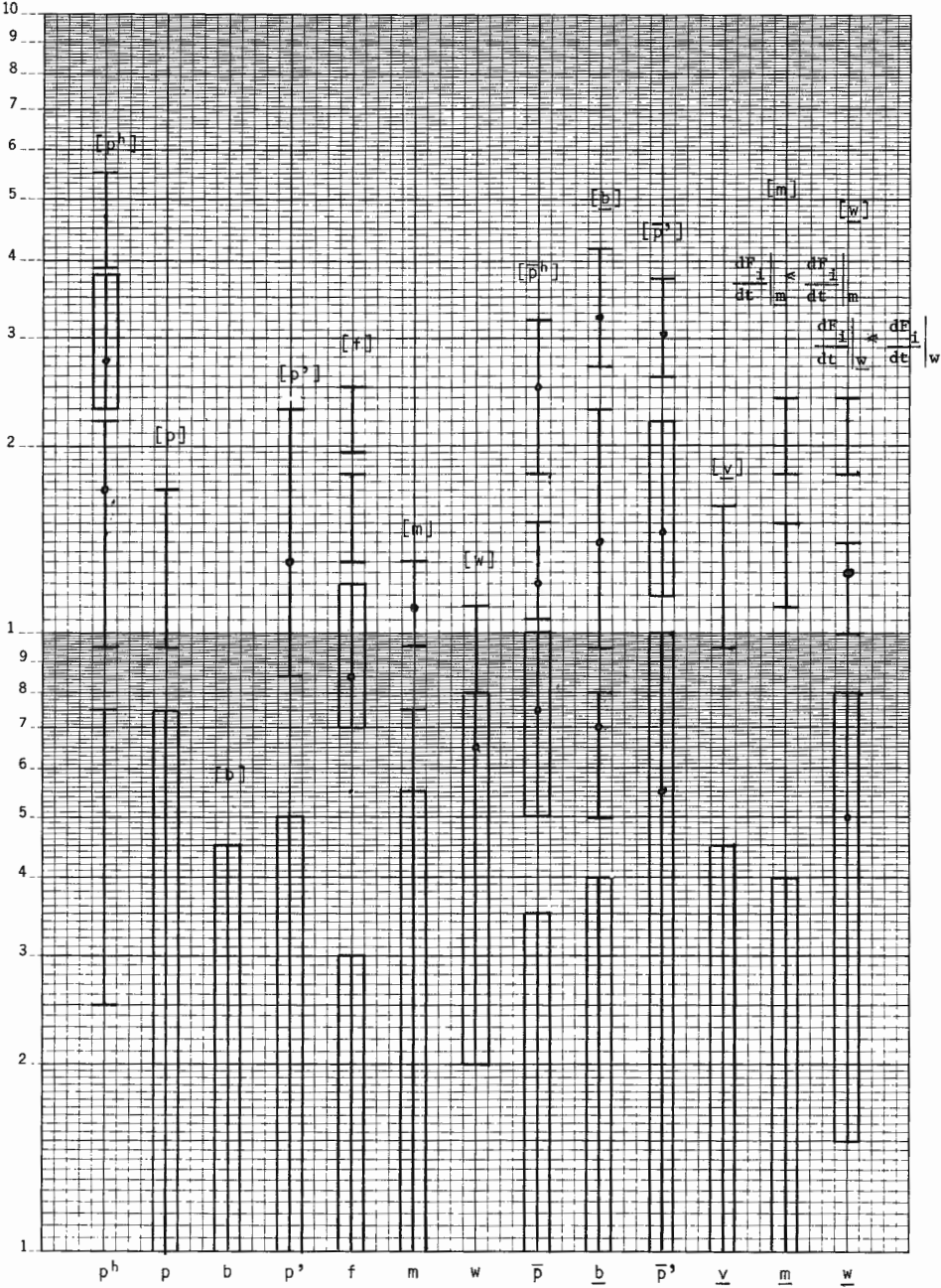
[+labial] (labio-dental spirants also [-distributed])

/p/ Shapsegh and Bzhedukh West Circassian,

/v, f'/ East Circassian; Abkhaz (/f'/ only in Bzyb suddialects),

/w/ only in Ubykh dialects (Dumézil 1965:266-69).

3.1 Graph 1 shows the labial stops, labio-dental spirants, labial nasal, and rounded labial glide. No data were available for /v/, /f'/ or /w/. Interesting coupling effects between source features and overall stop burst ranges can be seen for the labial stops. All but /p^h/ (in Ubykh written /p/) show the expected (Cooper et al. 1952: 600) low frequency stop burst ranges. The behavior of /f/ shows little similarity with that found in an earlier study (Stevens 1960:42-43). /m/ shows a typical nasal



Graph 1 Labials

pattern and /w/ shows damping out of any sound below 0.2 kHz (kilo-Hertz, 1 Hertz = 1 vibration per second), an effect of lip-rounding.

3.2. The labials in all the languages are stable and exhibit little phonological activity, the sole rule being one that derives a /p/ or /b/ from a /w/ before voiceless or voiced consonants, respectively. All bilabials are what Catford (1977a:141-42) would term **endolabials**.

3.3. The long /w/ of a poorly attested Ubykh dialect appears to be the reflex of “normal” Ubykh pharyngealized \bar{w} .

4. Pharyngealized labials: / \bar{p} , \bar{b} , \bar{p}' , \bar{v} , \bar{m} , \bar{w} /

[+labial, +constricted pharynx]

Only in Ubykh (Vogt 1963).

4.1. Graph 1 exhibits the stop burst and spirant forms for this unusual series. Whatever their subjective impression, which is one of heaviness, darkness, even turbidity, these sounds are louder than their plain counterparts and exhibit, somewhat surprisingly, significant high frequency energy (above 1.5 kHz). / \bar{w} / and / \bar{w} / also show much slower formant transitions than in the nonpharyngealized sonorants. This is represented on the graph by noting the second derivative of the formants with respect to time, d^2F/dt^2 , (a measure of the rate at which the formants “bend” from one target to another) as being smaller for the pharyngealized sonorants than for the plain. This effect dominates in Ubykh dial. / \bar{w} /.

4.2. Phonologically, these segments are relatively stable in Ubykh. A few arise from forthright assimilatory processes, as with / $\bar{w}\bar{a}q'\bar{w}\bar{a}$ / ‘shepard, guard’ and / $\bar{a}-\bar{w}\bar{q}'\bar{w}\bar{a}$ / ‘the shepard, guard’, with secondary / \bar{w} / (Dumézil 1965:248, entry 2091, though he denies this process on p. 209), or the variants / $\bar{c}a\bar{q}a\bar{p}\bar{a}$ / (Vogt 1963:98, entry 280).⁵

4.3. J. C. Catford (personal communication) has suggested to me that many of these unique segments have arisen through the influence of a contiguous lateral which must have been pharyngealized, as is still the case with the /l/ of (originally) nearby Abaza. Catford’s examples are convincing: Ub / $\bar{m}a\bar{s}'\bar{w}\bar{a}$ / ‘misery, hunger’, Abkh. / $\bar{a}-\bar{m}la\bar{s}'\bar{r}a$ / ‘famine’, Abaza / $\bar{m}la\bar{s}'\bar{y}ira$ / ‘idem’, Ub. / $\bar{b}\bar{y}\bar{a}$ / ‘thin, boney’, Abaza / $\bar{a}mlaga$ / ‘idem’, Ub. / $\bar{b}a\bar{q}'\bar{a}$ / ‘Caucasian fur hat’ ← */ $\bar{l}baq'\bar{a}$ / ← Turkish / $\bar{k}alpak$ / ‘idem’, Ub. / $\bar{w}\bar{a}$ / ‘dog’ ← */ $\bar{w}-la$ / grammatical class I-dog, Abaza / $\bar{l}a$ / ‘dog’, to which I might add Ub. / $\bar{y}\bar{a}$ -/ ‘to spy upon’, Circ. / $\bar{p}\bar{x}-a-$ / ‘look at’,⁷ and Ub. / $\bar{l}(\bar{a})\bar{w}\bar{a}$ / ‘excess, too much’, where the lateral seems still to be at work producing a pharyngeal. Suggestions that these pharyngeals are merely expressive (ideophonic) (Dumézil 1974:24, §9; 1975:144-45) cannot be seriously maintained in the face of comparative data. Dumézil’s best example in this vein, / $\bar{q}'a\bar{p}'\bar{a}wn\bar{a}$ / ‘handful-instrumental’ as opposed to / $\bar{q}'a\bar{p}'\bar{a}wn\bar{a}$ / ‘hand-instr.’, is subject to a clear explanation in that the

first word's sense allows an etymology involving either /b̥a-/ 'to pile up' or /b̥ə/ 'big, thick' (themselves likely related), so that the pharyngealized root likely goes back to something like */q'a-p'a-b̥(a/ə)-/ 'arm-tip-pile/big amount'. A few forms, however, remain unexplained, as with the just cited /b̥a-/ and /b̥ə/ and such forms as the personal name /p̥'əp'əz̥w/ (Esenç's maternal grandfather) (Vogt 1963:236)

4.4. The precise nature of the pharyngealization in this series will be taken up with the pharyngeals below §32.2.

5. Rounded labials: /b^w, p^w, m^w, f^w/

[+labial, +round] Only in Shapsegh West Circassian,

5.1. These segments were not available for spectrographic analysis.

5.2. The only one of these that is clearly a single segment is /f^w/, cf., the one morpheme Shapsegh /-f^w/, Chemgwi WC /-šχ^w/, Kabardian EC /-šyχ^w/, all 'big.' This treatment of this cluster in Shapsegh suggests that all sequences of /CW/ might be interpreted as /C^w/ in this dialect. Thus, several words with /p^w/ may occur, such as /p^wa/ 'place', and /p^waaʎa/ 'time, term' (Kuipers 1975:13), one with /b^w/, /b^wa-w/ 'very-adv(erb suffix)', and one with /m^w/, /m^wa-w/ 'here, now-adv'.⁸ No evidence for /p^{hw}/ or /p^w/ exists. Kuipers (1975:13; 1963:65) posits /p^w/ (and /t^w/) for Bzhedukh, but both of my Bzhedukh informants strenuously reject this interpretation and insist (despite native orthographic conventions) that this is a cluster, as is also the case for /bwa-w/ and /mwa-w/. Since Shapsegh distinguishes such pairs as /p'aast^ha/ 'mush' (with brief glottal hiatus) and /p'ʔa/ 'bed, mattress' (with long hiatus), or /s'ə-/ 'to make, build' and /s'ʔa-/ 'to know (trans.), to perform (intrans.)' (Kuipers 1975:13, 29), it may well have /p'ʔ^w/ instead of /p^w/ . Such a cluster would remove some of the analogical ground from under /b^w/ and /m^w/, so that these might be underlying clusters as well. This would leave only /f^w/, which because of this restriction to final position in the overwhelming number of icons would resist reanalysis as /fw/, i.e., /-f^w#/ fits into a dominant Northwest Caucasian pattern whereas /-fw#/ does not).

5.3. A spectrographic rendering of /p'ʔ^w/ can be found in Jakobson, Halle, and Fant (1951:50, fig. 6).

6. Denti-dental spirant: [T]

[+anterior, -labial, -coronal]

Shapsegh and Bzhedukh West Circassian only.

6.1. This extremely exotic sound consists of blowing air between clenched teeth. For want of any accepted symbol I have chosen [T]. Catford (1977a:148) notes it in a voiceless form as an allophone of the voiceless palatal spirant in Shapsegh. I have heard it in both Shapsegh and Bzhedukh, but only in casual, rapid speech. It never occurs in

elicitation forms and to date I have failed to record it in text reading. Regrettably, therefore, it has not undergone spectrographic analysis.

7. Plain dental: /t^h, t, d, t', n/

[+coronal, +anterior, -continuant]

/t/ in Shapsegh and Bzhedukh West Circassian.

7.1. Graph 2 shows the stop bursts for the dental stops. /n/ is the rightmost bar and, of course, has no stop burst. Rather /n/ exhibits a sonorant nasal pattern very much in agreement with that found early on for such sounds (Tarnóczy 1948:73, figure 2). The bursts for the dental stops are similar to those for the corresponding alveolars (Cooper et al. 1952:600). One must note, however, that while alveolar bursts begin around 2.5 kHz, dentals begin at 1.4. This distinct lowness is perhaps supplemented by a weaker stop burst for dentals than for alveolars (Jongman, Blumstein, and Lahiri 1985).

7.2. As with the labials, there is some coupling between stop burst and source feature. This is not merely a matter of voiced sounds having less energy and so showing a narrower frequency range on the spectrogram, or of ejectives having high energy and so a wider range, but rather of aspiration, at least, distinctly shifting the stop burst range higher and adding preponderant high frequency energy. This is an interesting effect and goes against the general position that stop burst parameters are entirely independent from other acoustic parameters (Oden 1978). Some complex coupling effects must be taken into account (Santerre and Suen 1981).

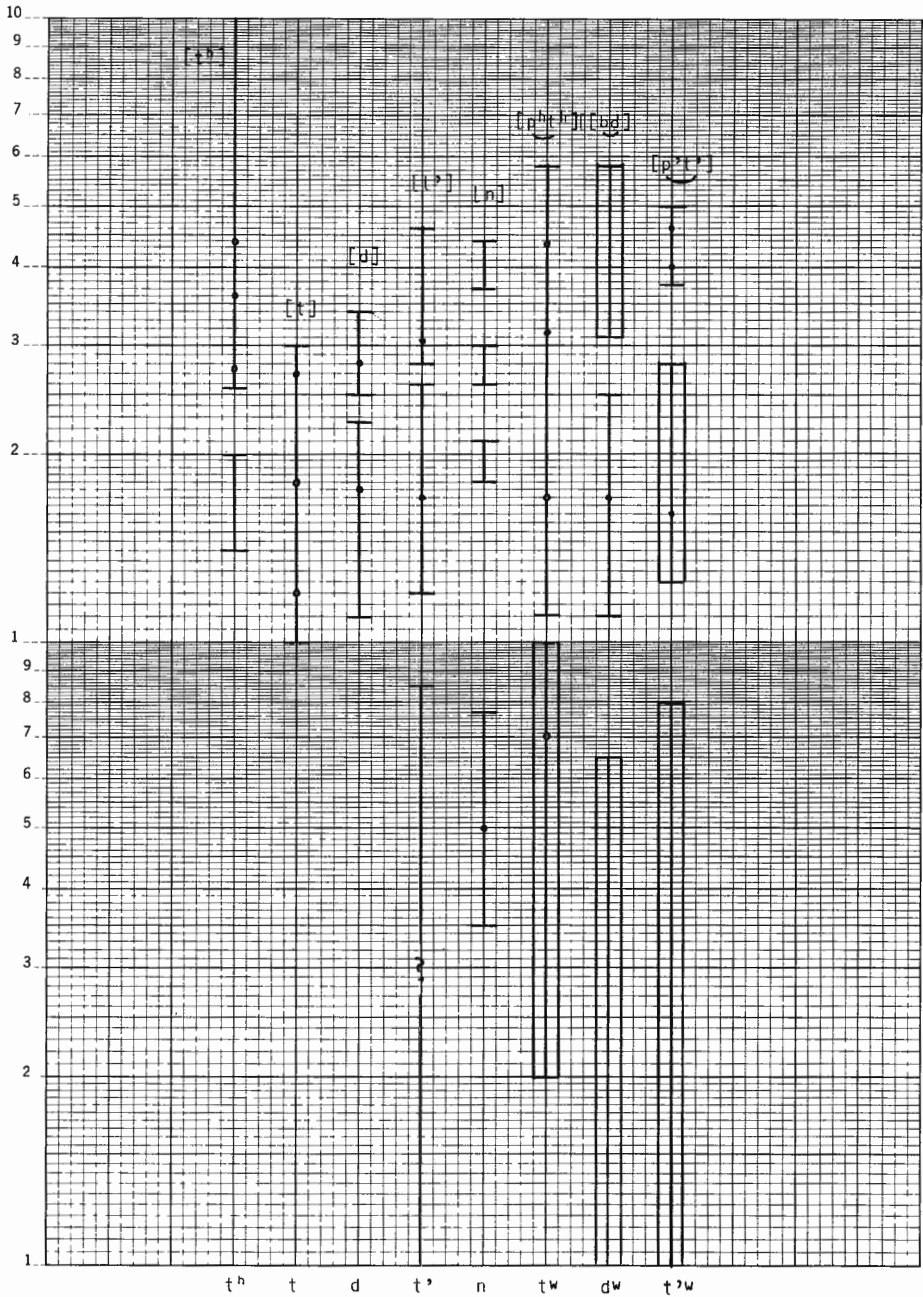
8. Rounded dental stops (non-affricated): /t^w, d^w, t'^w/

[+dental, +round]

Ubykh, Abkhaz dialects.

8.1. Graph 2 shows the stop bursts for these unusual sounds, wherein labial and dental bursts are seemingly superimposed. In fact these segments are realized with simultaneous labial and dental closures (Catford 1977a:146; 1977b:290; 1972:681). The labial closure exhibits a slight pouting and is what Catford (1977a:146) calls *exolabial*. Thus, graph 2 actually shows the stop bursts of [p^ht^h, b^d, p't']. Ubykh [b^d, p't'] and Abkhaz [p't'] show a tendency toward bilabial trilling at release, a rare instance of oralic egressive airstream mechanism.

8.3. Ubykh and Abkhaz appear to have shared a rule whereby coronal segments that were anterior and nonhigh had their rounding, if any, modified so that it agreed in manner with the primary closure. Under certain conditions these segments may lose their rounding and yield up simple dental stops: Ub. /-t^w-/ 'to be (sg.)', but /-t-/ 'being' as in /maa-tə/ 'where (he) is' (Vogt 1963:194, entry 1915); Bzyb Abkhaz (Kaladakhwar sub-dialect), /-t'^wa-/ 'to sit', but /w-t'ä/ you (male)-sit, 'sit!' (said to a man),



Graph 2 Dental Stops

9. Alveolar trill and laterals: /r, l, ɭ, ʎ, λ, λ'/

9. Alveolar trill and laterals: /r, l, ɭ, ʎ, λ, λ'/

/l/ Abkhaz, Abaza, Ubykh and Circassian when under strong Turkish or Arabic influences.

/t/ Ubykh and Circassian when under strong Turkish or Arabic influences.

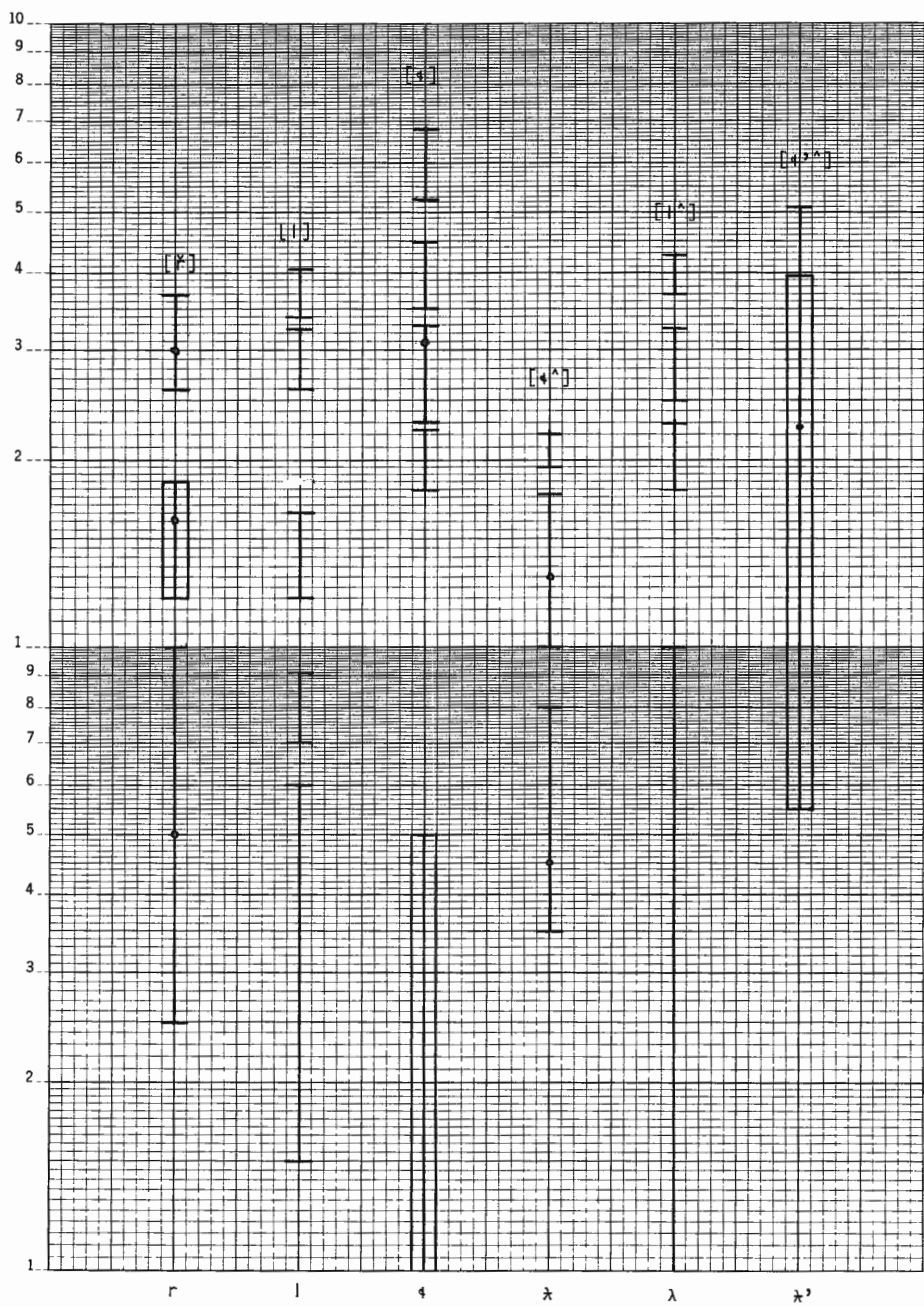
/λ, λ, λ'/ Circassian and Kabardian loans in Abaza; Ubykh also has /λ'/.

9.1. The alveolar trill and the various laterals are plotted in graph 3. /r/ is invariably an alveolar trill. It has a clear formant pattern and closely resembles the trilled-*r* studied by Tarnóczy (1948:73, figure 2).

9.2. The laterals are a confusing matter. The /l/ in the graph is from Ubykh. Ubykh has been strongly influenced by Turkish for over a century, so that this liquid lateral now reflects both an original */l/, which Ubykh seems at one time to have shared with Ahkhaz-Abaza, and */λ/, a non-sonorant lateral spirant now confined to “pure” forms of Circassian. Ubykh /l/ is light and its formant patterns agree fairly well with the light *l* of Tarnóczy’s 1948 study. Abaza has a dark-*l* (Catford, personal communication). The Abaza tapes which I received from W. S. Allen (through Calvert Watkins) seem also to confirm this, but their fidelity is too poor to permit spectrographic analysis. Ubykh /l/ may also be dark under some conditions (cf. §4.3, above). One might expect a velarized or pharyngealized /l̥/ to lack energy between 2.0 and 3.0 kHz, as is the case with Tarnóczy’s dark-*l* (ibid.).

9.3. Ubykh /ɬ/ shows a pseudo-formant (formants composed of aperiodic sound) structure differing both from the formant structure of /l/ and the pseudo-formant structures of either Circassian /χ/ or /λ/, both of which are spirants. In other words, Ubykh /ɬ/ is not merely a voiceless lateral spirant. One might better term it a voiceless sonorant, taking [+sonorant] as based upon degree of oral occlusion (Chomsky and Halle 1968:302) and not upon quality of voicing, if any. The differences between /l/ and /ɬ/, therefore, are most likely a pronounced example of the coupling between sources feature and point of articulation which has been already shown for /p^h, t^h, t^w/.

9.4. The lateral spirants of the Circassian languages are most unusual. They are alveolars, but made with extreme occlusion. Only / λ' / has a tendency toward affrication. Graph 3 shows only Circassian segments, Ubykh / λ' / having been omitted. / λ / and / λ' / show pseudo-formant structures, but neither resembles each other nor any of the other laterals very closely. The most remarkable sound is the ejective / λ' /, which



Graph 3 Coronal Sonorants, Lateral Spirants

consists of a mass of high energy noise running from 0.55 to 3.95 kHz, tapering off at 5.1 kHz. No structure is discernable on the spectrogram, although a frequency vs. energy plot might show some maxima. As with the other ejectives, one can safely assume that increased pressure has caused a boost in this segment's overall energy, but this does not explain the bottom cutoff at 0.55 kHz when the next highest bottom cutoff, that for / λ /, is only at 0.35 kHz.

9.5. If one is committed to a quantal theory of speech perception (natural articulatory groups are characterized by distinctive, shared regions of acoustic activity), then the laterals as presented by Ubykh and Bzhedukh present some major problems. One might seek to explain some of the variation in graph 3 as due to coupling between various source features and sonorant or spirant supraglottal configurations, but this seems highly unpromising when one notes that acoustically /l, ɭ, λ /, two sonorants and a spirant, one voiced and two voiceless, form a natural group, whereas / λ , λ' / stand apart from the rest as well as from each other.

10. Alveolar affricates and spirants, plain laminal: /c^h, c, z, c', s, z, s'/

[+coronal, +anterior, \pm delayed release, \mp continuant, +distributed]

/s'/ Shapsegh West Circassian and Terek Kabardian (Abitov et al., 1957:21)

10.1. Graph 4 presents the noise ranges of the various coronal spirants, starting with this series. The affricates have not been plotted, though as a rule their stop bursts are lost in the noise of their spirant release so that they look precisely like short spirants. There is also coupling between source feature and spirant noise, the chief effects being the reduced noise output of voiced forms and the enhanced output of aspirated and ejective ones. I have omitted voiced forms, but have plotted the more exotic ejectives (§14, 15) and aspirated spirants (§16, 18).

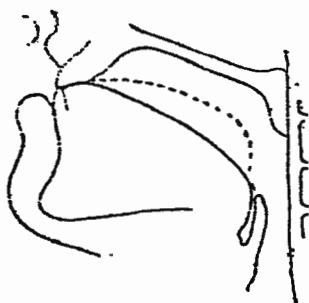


Figure 1 Bzyb Abkhaz /c'/

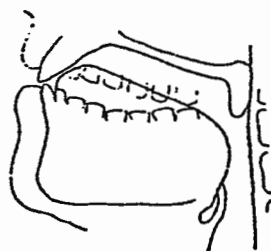
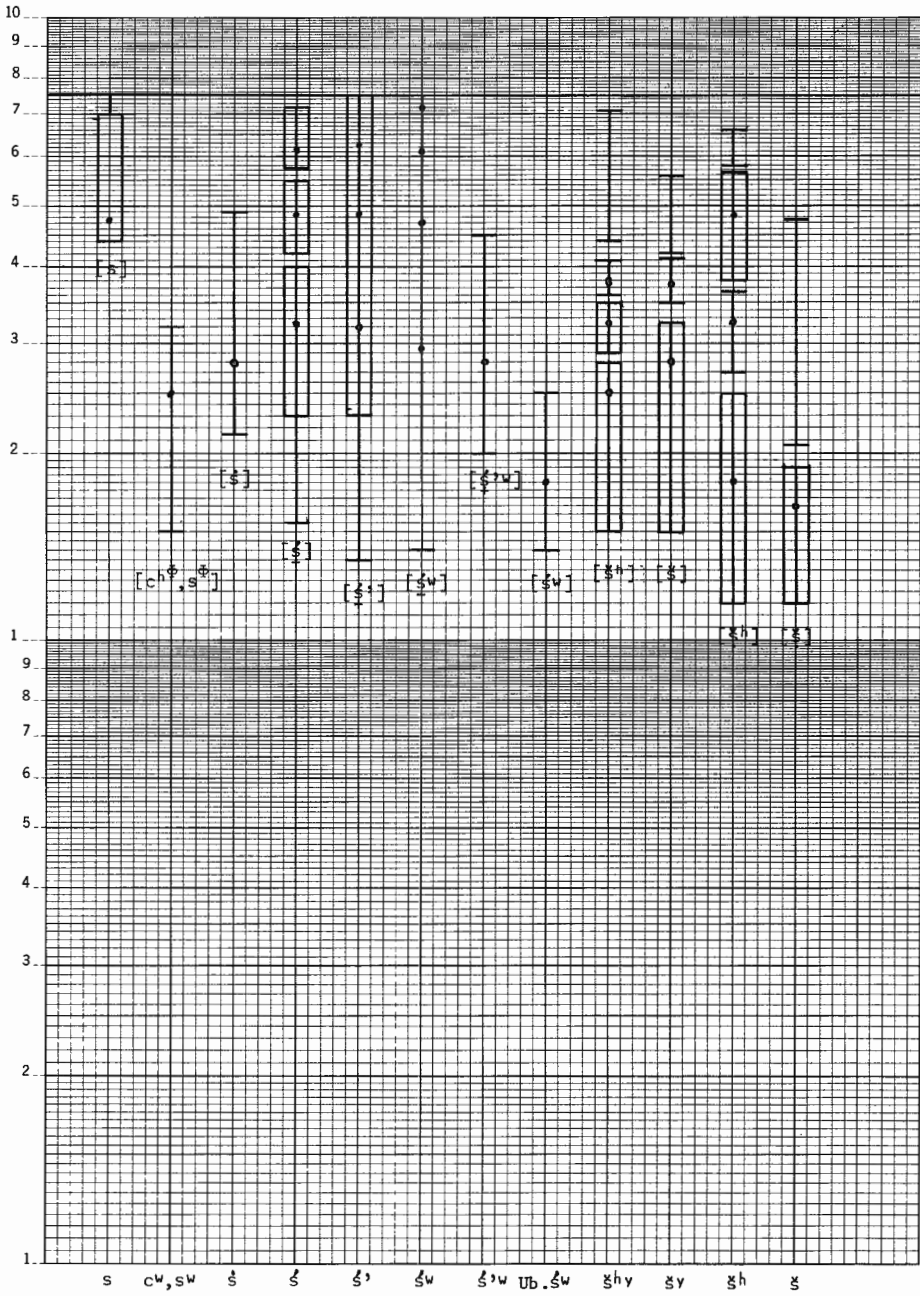


Figure 2. Bzyb Abkhaz /s/



Graph 4 Coronal Spirants

10.2. Figure 1 shows Bzyb Abkhaz /c'/, figure 2 /s/ (Bgažba 1964:88, 78), figure 3 Ubykh /s/ (Leroy and Paris 1974:264).



Figure 3. Ubykh /s/

These figures will serve as useful comparisons later when the apical coronals are examined (§§12, 13). The Ubykh tracing (of Tevfik Eseng) looks a bit odd, but a laminal coronal appears to be represented.¹⁰ The dashed line in this and the other Bzyb tracings to follow represents the highest portion of the outer margins of the tongue.

10.3. The Bzyb data was not available for spectrographic analysis. The graph represents several icons of Ubykh and Bzhedukh /s/, both very similar acoustically. There is nothing odd here, the noise beginning at 4.4 and running to 7.0 kHz, still to be seen, tapering off from 7.0 to 7.5 kHz, the limit of the machine. The laminal alveolars have the highest frequency noise of any segment.

11. Alveolar affricates and spirants, rounded laminal: /c^w, ʒ^w, c^ʷ, s^w, s^{·w}, z^w/

[+coronal, +anterior, ±delayed release, ∓continuant, +distributed, +round]

Only in Ubykh; (/s^{·w}/ only in Ubykh dialect (Dumézil 1965:266-69)).

11.1. The phonological rule of §8 applies here to produce bilabialized affricates and spirants: [c^β, ʒ^β, c^{ʰβ}, s^β, z^β]. Vogt (1960) is very clear about this bilabialization, saying that it is plainly visible in the course of the articulation. Nevertheless, almost anyone who has at all worked on Ubykh comes up with a different interpretation of this series (cf. Leroy and Paris 1974 for a survey of various analyses of Ubykh). The phonological rule which governs the realization of these segments, as well as those in §8 above and §13 below, is shown in (1).

(1) Rule Governing Labialization of Rounded Dental/Alveolar Series in Ubykh and Abkhaz

$\begin{array}{c} [+round] \rightarrow \\ 2 \qquad \qquad 2 \end{array}$	$\begin{array}{c} -round \\ +labial \\ \alpha\text{continuant} \\ \beta\text{distributed} \\ \langle +continuant \rangle \end{array}$	/	$\begin{array}{c} 2 \qquad \qquad \qquad 2 \\ \hline +coronal \\ +anterior \\ -high \\ \alpha\text{continuant} \\ \beta\text{distributed} \\ \langle +delayed release \rangle \end{array}$
	2		1

What (1) does is to modify the secondary feature of rounding so that it becomes a labialization matching the primary closure in continuity and distributedness, with the proviso that if an affricate is being dealt with, its continuant release dominates the whole segment and produces a labial continuant throughout. In terms of internally structured segments (of which affricates may be considered an example), rule (1) might be rendered as (2) (Halle and Clements 1983:14).

(2) Variant of (1) Using Internally Structured Segments

$\begin{array}{c} [+round] \rightarrow \\ 2 \qquad \qquad 2 \end{array}$	$\begin{array}{c} -round \\ +labial \\ \alpha\text{continuant} \\ \beta\text{distributed} \\ \langle +continuant \rangle \end{array}$	/	$\begin{array}{c} 2 \qquad \qquad \qquad 2 \\ \hline +coronal \\ +anterior \\ -high \\ \alpha\text{continuant} \\ \beta\text{distributed} \\ \langle -continuant \mid +continuant \rangle \end{array}$
	2		1

The important thing is the linking of the distributed quality of the primary articulation with that of the secondary. This series is laminal in Ubykh and therefore distributed. The resulting labialization from (1)/(2) is also distributed and therefore bilabial. In §13 I shall show that the analogous series in Abkhaz (which is a complete series only in the Bzyb dialect) is apical and therefore nondistributed, viz., labiodentalization. Ubykh and Abkhaz provide the only examples that conclusively show [distributed] to be a true phonological feature, abstract and transcending any particular articulation (Colarusso 1979:307-8). Any account of the Ubykh series that differs from the present one in a substantive way is wrong both phonetically and phonologically.

11.2. Figure 4 is taken from Leroy and Paris (1974:265) and is supposed to represent Ubykh /s^w/. This tracing has caused much confusion and wonder for it clearly shows a labialized palatal spirant, rendered by most workers in Caucasology by /χ^w/, (which by IPA standards might be misconstrued as a rounded uvular). In fact the authors have discovered a hitherto unrecognized phoneme of Ubykh. To see this one must step outside of Ubykh to some of its sister languages.



Figure 4. Ubykh supposed /s^w/.

11.3. In West Circassian, Proto-Circassian */χ^w/ has been realized as /f/ except in the cluster */šχ^w/ which has undergone its own development (Kuipers 1975:61-63, where /x^w/ is used for /χ^w/). This suggests that at one time the rounding of */χ^w/ must have been subject to a rule such as (1)/(2). In Abaza this is in fact still the case. Abaza has /χ^w/ in Kabardian borrowings and this is realized as [χ^φ] (Allen 1956:129, table 1). As Ubykh was originally just over the Caucasian massif to the west of Abaza, it may have shared in an old areal feature by realizing /χ^w/ in this way (perhaps palatals are [+distributed]). In fact, all the forms used to make the x-ray tracing in question, /s^wa/ 'white', /s^wà-s^wa/ lit. 'sea-white', i.e. 'Aegean Sea', /məs^wà/ 'day', /s^wəq'à/ 'oil, grease', (Leroy and Paris 1974:262), all have cognates in Kabardian that contain a /χ^w/, as may be seen in (3).

(3) Ubykh-Kabardian Cognates

	<u>Kabardian</u>	<u>Ubykh citation</u>	<u>Ubykh actual</u>
(i) 'white'	/χ ^w ə/	"s ^w a"	/χ ^w à/
(ii) 'oil, grease'	/(šə)χ ^w a/	"s ^w əq'à"	/χ ^w əq'à/
(iii) 'day'	/maax ^w a/	"məs ^w a"	/məχ ^w à/

In fact the entire Ubykh corpus must be reviewed and revised in the light of comparative data to sort out the forms with true /s^w/ from those with /χ^w/. This is not hard to do (Colarusso, 1992), so that one can safely say

that had Leroy and Paris examined forms such as those in (4), little confusion surrounding the rounded alveolar affricates and spirants would have arisen.

(4) Ubykh Forms with True /s^w/

	<u>Ubykh</u>	<u>Bzhedukh West Circassian</u>
(i)	/s ^w əɣ ^w à(ɬa)/ 'you (plural)'	/š ^w a/ 'idem'
(ii)	/s ^w aaš ^y à/ 'shirt'	/š ^w aaša/ 'outfit, uniform'
(iii)	/s ^w àχ/ 'smooth, even, flat'	/š ^w afə/ 'plain, steppe'

The forms in (4) are taken from Vogt (1963) and Kuipers (1975). List (4) could be greatly extended, as could List (3).

11.4. If Leroy and Paris had made use of the voiced spirant of this series or some of the affricates, they would not have produced the odd tracing in figure 4. In fact they filmed /z^wà/ 'sky', /a-z^wà/ 'the-crowd',¹¹ and /a-z^wa-z^wà/ lit. 'the-drink-drink', i.e. 'wedding reception, party in general,' but seem not to have used these in making x-ray tracings, confining themselves instead to putative "s^w." I have been careful in my spectrographic work to rely either upon affricates of this series or upon forms with actual /s^w/. The results are plotted as the second column from the left in graph 4. The noise for this series is significantly lower than that for its plain counterpart, bottom cutoff of 1.5 kHz as opposed to 4.4, and is damped out above 3.2 kHz. In §15 I shall show that rounding has effects similar to this labialization. It is noteworthy that despite similarities in articulation the plain alveolar affricates and spirants are acoustically disjointed from their labialized counterparts.

12. Alveolar affricates and spirants: plain apical /č, ž, č', ž'/

[+coronal, +anterior, ±delayed release, ∓continuant, -distributed]
Only in Ubykh, Bzyb, Abkhaz and Abaza (Lookət dialect of Anatolia).

12.1. Figures 5 and 6 show Bzyb /č'/ and /š/, respectively (Bgažba 1964:87, 72), and figure 7 shows Ubykh /š/ (Leroy and Paris 1974:264). Again the Ubykh tracing appears a bit odd. The Bzyb ones clearly show apical articulations, however, and the Ubykh tracing resembles them more closely if one compares it with figure 3, Ubykh /s/, which latter is clearly laminal. The third column from the left in graph 4 shows Ubykh /š/ with energy running from 2.15 to 4.9 kHz and a center about 2.8. The overall effect is to lower the usual alveolar target and to damp out high frequency energy, but not as much as in labialization.

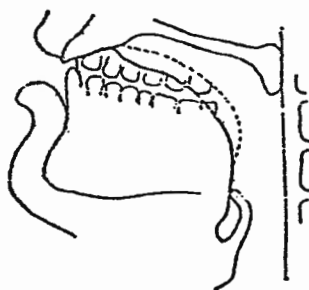


Figure 5. Bzyb Abkhaz /č'/



Figure 6. Bzyb Abkhaz /š/



Figure 7. Ubykh /š/

12.2. The Abaza dialect with this series (Dumézil 1970:276) has merged /č^w, ž^w, č^{'w}, š^w, ž^w/ with an original */č^w, ž^w, č^{'w}, š^w, ž^w/, still found weakly labiodentalized in the Lo and Kuban dialects of Abaza (Genko 1955:39, 48), and then to have unrounded them to produce /č, ž, č', š, ž/. Lookət Abaza is then the only dialect other than Kabardian to lack any rounded coronals whatsoever.

13. Alveolar affricates and spirants, rounded apical: /č^w, ž^w, č^{'w}, š^w, ž^w/ [+coronal, +anterior, ±delayed release, ∓continuant, -distributed, +round]

Bzyb (Bgažba 1964), Ashkharawa Abkhaz (Lomtatiže 1954), and Lo-Kuban Abaza,¹² /č^w, ž^w, č^{'w}/ alone in all other Abkhaz dialects.

13.1. Figure 8 shows Bzyb Abkhaz /ž/ which is to be compared to figure 9, with Bzyb /ž^w/ (Bgažba 1964:76). One can plainly see that /ž^w/ is both apical and labiodentalized. This series, therefore, is subject to a rule of the form (1)/(2). The only difference between this series and Ubykh /c^w, ž^w, c^{'w}, s^w, z^w/ is that the former is [-distributed] whereas the latter is [+distributed], (cf., §11). Unfortunately, no data were available for spectrographic analysis.



Figure 8. Bzyb Abkhaz /ʒ/

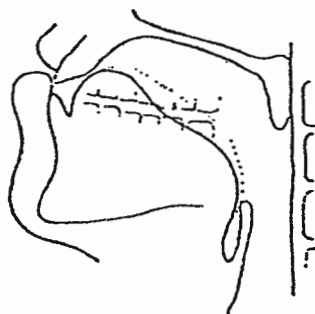


Figure 9. Bzyb Abkhaz /ʒʷ/

13.2. The labiodentalized spirants are opposed to rounded alveolopalatal spirants (§15) only in Bzyb, Ashkharawa Abkhaz and Lo-Kuban Abaza. In the other Abkhaz dialects the affricates have remained while their spirant counterparts have merged into the rounded alveolopalatals, i.e., Bzyb has /čʷ, ʒʷ, čʷ, śʷ, ʒʷ/ and /śʷ, ʒʷ/, while all others have only /čʷ, ʒʷ, čʷ/ and /śʷ, ʒʷ/. The rounded affricates are realized as [čʰ, ʒʰ, čʰ] in the other Abkhaz dialects, while /śʷ, ʒʷ/ are always [śʷ, ʒʷ] (Catford 1972:680, but see §16 for further discussion of this pair). Hewitt (1979:256) makes the point that the labiodentalization of the affricates is somewhat lopsided, with the right side of the lower lip held more tightly against the teeth and the left side looser or even bowed out slightly. Wim Lucassen (personal communication) has confirmed this odd lateralized labiodentalization. Despite skewing of the series, however, there is little reason to lump /čʷ, ʒʷ, čʷ/ together with /śʷ, ʒʷ/, as most workers have done, since these two series behave distinctly with regard to rule (1)/(2), both phonetically and phonologically (see §16 for further discussion of /śʷ, ʒʷ/).

13.3. Labiodentalized sounds occur within the Caucasus in the Northeast Caucasian languages Tabasaran, Lezgi and Aghul (Catford 1972:681) and outside the Caucasus in the African languages Kom and Kutep (Ladefoged 1864:31-32, plate 16A). In Kutep labiodentalization has the effect of lowering the primary target and adding spirant noise characteristics of [f] or [v], analogous to the superimposition of labial bursts onto dental ones in [pʰᵗʰ, bᵈ, pʰᵗʰ]. One might therefore expect the acoustics of [čʰ, ʒʰ, čʰ, śʰ, ʒʰ] to be a lowered version of the plain apical alveolars with labiodental spirant noise super-imposed.

14. Alveolopalatal spirants with advanced tongue root: /ś, ʒ, śʰ/

[+coronal, +anterior, +high, +ATR, +distributed]

Only in Circassian languages.

14.1. This series is made at a point of articulation which is included in no phonetic text that I have examined. The sounds are made with the blade of the tongue running from the alveolar ridge to a point midway up on the hard palate, the tongue tip being pressed against the lower teeth (Kuipers 1960:20; Catford 1942:16). Kuipers and Catford have both suggested that some velarization or pharyngealization might occur in this series, but one of the few instrumental studies of Kabardian (Henderson 1970:95) shows no evidence for this. Israeli Shapsegh, however, appears to have pharyngealized this series [+CP], /s, z, sʔ/ and to have made it alveolar, but this may be viewed as a secondary and recent effect due to Arabic influence. These segments appear to be made instead with advancing of the tongue root ([ATR]) (Perkell 1971; Halle and Stevens 1969; Pike 1967; Stewart 1967), sometimes called the “expanded pharynx.” This advancement accounts not only for the contact between tongue tip and lower teeth (due to pronounced advancement of the whole tongue), but also for the lowering effect upon the first formant of an adjacent vowel. This lowering effect mimics the behavior of velarization and pharyngealization, and thus accounts for Kuipers’ and Catford’s earlier impressions. The Armavir dialect of West Circassian (a subdialect of Bzhedukh) (Bouda 1939) and the Terek and Baksan dialects of Kabardian (Kuaševa 1969; Mamrešev 1969) have lost [ATR] to produce palatalized “s”: [+coronal, +anterior, +high], /sʲ, zʲ, sʲʔ/.

14.2. The only segments available for analysis were Bzhedukh /ś, ź, śʔ/. The fourth and fifth columns from the left in graph 4 show /ś/ and /śʔ/, respectively. /ź/ is like /ś/ save that its overall energy is lower and its resulting spectrogram tracing consequently narrower. The ejective spirant, on the other hand, has an enormous amount of energy because of its greater supra-glottal pressure. As a result its spectrogram looks wider than that of the simple voiceless spirant. The range overlaps with that of /s/, but extends down to a much lower cutoff, at least below 1.6 kHz. Despite the overlap between this series and the others, its overall acoustic signature is distinctive.

14.3. The ejective spirant is characterized not merely by greater energy, but by creaky voice in Eastern Circassian dialects (Kuipers 1960:19) or by a glottal hiatus, which is very pronounced in Bzhedukh (Catford 1977a:19), roughly 100 ms., leading to the neutralization of the contrast between /śʔ/ and /śʔʔ/, which is retained in Shapsegh where the glottal hiatus is not so great.

14.4. Some affricates have been claimed for this series (/ǰ/) for Chemgwi (Jakovlev 1930), and /č/ and /ǰ/ for Abadzakh (ibid.), but the only affricates I have encountered despite extensive work are secondary

ones that occasionally emerge in Bzhedukh, e.g., /s-šha/ → /čha/ 'my-head,' in about half the icons.

15. Alveolopalatal affricates and spirants with [+ATR] and rounded:
/č^{hw}, č^w, ž^w, č^w, š^w, ž^w, š^w/

[+coronal, +anterior, +high, +ATR, +distributed, +round]

In West Circassian.

For some speakers of Tapanta Abaza (without /š^w/ (Genko 1955:52).

/č^w/ only in Kuban sub-dialect of Shapsegh West Circassian (Keraševa 1957:36).

15.1. Graph 4 exhibits the acoustics of /š^w/ and /ž^w/ in columns 6 and 7 from the left, respectively. They are very similar to their nonrounded counterparts save that their energy is less and their lowest maxima are below 3.0 kHz, whereas those for /š/ and /ž/ are above 3.0 kHz. The most obvious and perhaps the crucial difference is a damping-out of higher frequency noise due to the effects of rounding. This is particularly evident with the ejective. As with the nonrounded segments the ejective has more energy in its range of 2.0 to 4.5 kHz than does the voiceless segment.

15.2. A similar series occurs for some speakers of the Tapanta dialect of Abaza, perhaps due to West Circassian influence but nevertheless in native vocabulary. The West Circassian series is clearly [+ATR], as may be seen by its closeness to the non-rounded series. This is probably not the case with the Abaza series, though no firm data are available to confirm this conjecture.

16. Alveolopalatal spirants, rounded: /š^w, ž^w/

[+coronal, +anterior, +high, +distributed, +round, -ATR]

In Ubykh and all Abkhaz dialects save the Lyxny subdialects of Bzyb (Bgažba 1964:23).

16.1. Column 8 of graph 4 shows the acoustics of Ubykh /š^w/. Like its Circassian counterpart, the Ubykh sound starts at 1.4 kHz, but quickly damps out above 2.5 kHz, unlike the Circassian correlate. The Ubykh segment is not made with advancing of the tongue root as may be seen in figure 10 (Leroy and Paris 1974:265). The acoustic aspects of this minor articulation difference are marked, but this effect will be seen again in §23 where [-ATR] Circassian /χ/ is compared with [+ATR] Ubykh /χ/

16.2. The Ubykh forms show lip-rounding (Vogt, personal communication), which by virtue of their anterior point of articulation has an anterior quality: [š^u, ž^u]. The Abkhaz situation is a bit perplexing. For the southern dialects (Abzhwi-Samurzakan) Hewitt (1979:256, 258) speaks of labialization, [š^β, ž^β], while Catford (1972:680) mentions rounding for the Bzyb dialect to the north, [š^u, ž^u]. Later Catford

(1977a:190-91) mentions a palatal component for these segments, particularly in Bzyb, and uses symbols that suggest both labialization, [ɸ̥, β̥], and rounding, [ɸ̞, β̞]. It may well be that the Bzyb segments are [+ATR] and that the palatal component is due to this more advanced tongue position creating a greater area of occlusion. It may also well be that these segments are labialized in which case rule (1)/(2) has been generalized in Abkhaz (with the possible exception of Ashkharawa) to cover all coronals, making this series [+distributed] with [+distributed] labialization. Only further work will settle this confusing matter.

16.3. There are no affricates for this series.



Figure 10. Ubykh /šʷ/

17. Palatoalveolar affricates and spirants, laminal: /čʰy, čʷy, žʷy, čʷy, šʰy, šʷy, žʷy, šʷy/

[+coronal, -anterior, +high, +distributed, ±delayed release, ±continuant]

/čʷy/ only in Shapsegh and Bzhedukh West Circassian.

/šʰy/ only in Shapsegh, Bzhedukh and Abadzakh West Circassian (Balkarov 1970:137, 189).

/šʷy/ only in Abadzakh West Circassian (ibid.) and Kubano-Zelenchuk East Circassian (Bagov 1968).

Absent from all dialects of Kabardian East Circassian save for Baksan (Mamrešev 1969) and Malka (Šagirov 1969, 1961).

17.1. This series is made at the same point as English 'ch,' 'j,' etc., but it is customary among works dealing with this family to emphasize the laminal nature of the series with a small superscript 'y' diacritic.¹³ This is done because most often this series contrasts with a retroflexed or velarized palatoalveolar one.

17.2. Columns 9 and 10 of graph 4 show Bzhedukh /šʰy/ and Bzhedukh and Ubykh /šʷy/, respectively. Bzhedukh offers a true rarity in phonology, a contrast between aspirated and unaspirated voiceless spirants. Both are represented here, but the spectrogram does not reveal very clearly

the formant-like structure that has been imposed on the more usual spirant noise in /š^{hy}/. Another salient acoustic cue to aspiration in /š^{hy}/ is a brief transition between the spirant and any following noise which has the same sort of *h*-like quality that one would expect from any other aspirated segment.

17.3. Figure 11 shows the point of articulation for Ubykh /š^y/ (Leroy and Paris 1974:265) and figure 12 that for Bzyb Abkhaz /š^y/ (Bgažba 1964:80). The tongue body is raised and the corona is employed, so that they are [+coronal, +high]. The low frequency for the Bzhedukh and Ubykh segments is 1.5 kHz, roughly that for the alveolopalatals – plain or rounded, but there is more energy in this low range than there is for the alveolopalatals.



Figure 11. Ubykh /š^y/

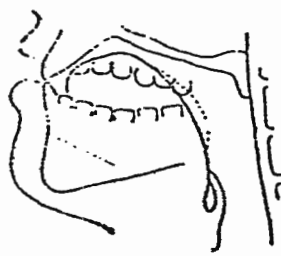


Figure 12. Bzyb Abkhaz /š^y/

18. Palatoalveolar affricates and spirants, laminal, rounded: /č^{yw}, ž^{yw}, č^{yw}, ž^{yw}, š^{yw}, ž^{yw}/

[+coronal, -anterior, +high, +distributed, ±delayed release, ±continuant, +round]

Dudaruko-Bibard, Lo-Kuban and in some Tapanta Abaza (Genko 1955:48; Allen 1965, 1956).

18.1. The series described in §17 appears to have a rounded counterpart in many dialects of Abaza. One would expect the acoustics of such a series to resemble closely that of the one in §17, but to show significant damping out of the upper end of the range. No data were available for analysis.

19. Palatoalveolar plain (abruptly released stops): [č^h, č, ʒ, ʈ]

[+coronal, -anterior, +high, -delayed release, -continuant]

Shapsegh West Circassian.

19.1. These simple palatoalveolar stops occur as allophones of the palatal stops /k^{hy}, k^y, g^y, k'^y/ in the Kirova subdialect of Shapsegh West Circassian (Keraševa 1957:30-31), and I have heard them frequently in other Shapsegh subdialects as well. Kabardian shows a similar tendency

to realize its palatal series with coronal segments (Kuipers 1960:26), but these are affricates in Kabardian. In Shapsegh these coronal allophones of the palatals are not affricated so that this dialect opposes a complete series of palatoalveolar affricates to one of plain stops, a most unusual situation. Plain palatoalveolar stops are quite rare, occurring elsewhere among Ugrian languages, such as Hungarian, and among a few Amerindian languages, such as some of the Mayan dialects. The available data has not yet been analyzed.

20. Palatoalveolar affricates and spirants, apical (retroflexed): /č^h, č, ž, č', š^h, š, ž, š'/

[+coronal, -anterior, -high, -distributed, ±delayed release, ∓continuant]

In West Circassian, except Shapsegh and Hakuchi dialects.

/š, ž/ only in Kuban Kabardian (Kumakhov 1969, 1961), otherwise absent from this language, as well as from Bzyb Abkhaz.

/š^h/ in Bzhedukh and Abadzakh, /š'/ only in Abadzakh West Circassian.

20.1. The last two columns (11 and 12) on the right of graph 4 show the acoustics of Bzhedukh /š^h/, and Bzhedukh and Ubykh /š/. They show low cutoffs at 1.15 kHz and a general lowering to the acoustic center of gravity when compared with their laminal counterparts. Figure 13 shows Ubykh /š/ (Leroy and Paris 1974:266) and this shows strong retroflexion against the prepalatal or palatoalveolar zone. The same is true of the Bzhedukh counterparts of this series and seems to hold true for the Abkhaz dialects other than Bzyb (Hewitt 1979:256-58). Here again the formant-like pattern of the unusual voiceless aspirated spirant of Bzhedukh is not revealed very well by the spectrogram.



Figure 13. Ubykh /š/

21. Palatoalveolar affricates and spirants, apical (retroflexed), rounded:/č^w, ž^w, č^w, š^w, ž^w/

[+coronal, -anterior, -high, -distributed, ±delayed release, ∓continuant, +round]

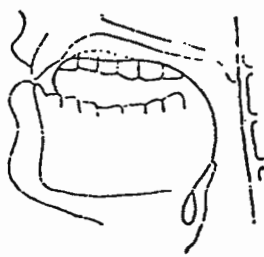
Kum-Lo subdialect of Abaza (Genko 1955:48).

21.1. These occur as the rounded coronals in only one dialect complex of Abaza. The acoustics of this series is likely to resemble that of its unrounded counterparts (§20), but to exhibit damping off higher frequency noise. Unfortunately no data were available for analysis.¹⁴

22. Palatoalveolar affricates and spirants, laminal, velarized: /č⁻, ž⁻, č^{'-}, š⁻, ž⁻/

[+coronal, -anterior, +high, +back, +distributed, ±delayed release, ∓continuant]

Only in Bzyb Abkhaz (Bgažba 1964).

**Figure 14. Bzyb Abkhaz /š/**

22.1. Figure 14 shows Bzyb Abkhaz /š/ (usually recorded in Western works simply as /š/ and in Soviet works as /šə/, as with the retroflexed series) (Bgažba 1964:80). One might even wish to argue that this series was a velarized, apicoalveolar one if it were not for the outer margins of the tongue blade (the dotted line) coming into approximation with the prepalatal area. Some subtle contrasts between this type of segment and sequences of palatoalveolar plus uvular spirant can result, as in (3), (taken from Bgažba 1964:80).

(3) Bzyb Abkhaz Contrasts Involving Velarized Palatoalveolars

- | | |
|--------------------------------|------------------|
| (i) /a-š ⁻ à/ | the-cord, rope |
| (ii) /a-š ^y àč/ | the-grass, sword |
| (iii) /a-š ^y xaran/ | the-queen bee |
| (iv) /à-š ⁻ x/ | the-mole cricket |

This articulation appears to be unique to Bzyb. No data were available for analysis.

23. Palatal stops and spirants: /k^{hy}, k^y, g^y, k^y, χ^h, χ, ğ/

[+high, -back, -coronal, -anterior, ±continuant]

/k^y/ only in Shapsegh and Hakuchi.

/χ^h/ only in Shapsegh and Bzhedukh (Rogava and Keraševa 1966:table 2 facing p. 56, f.n. 7).

/χ/ absent from Abkhaz save in Circassian loans into Ashkharawa

/χ/ in Circassian loans in Abaza, and perhaps in two or three native words (Catford, personal communication).

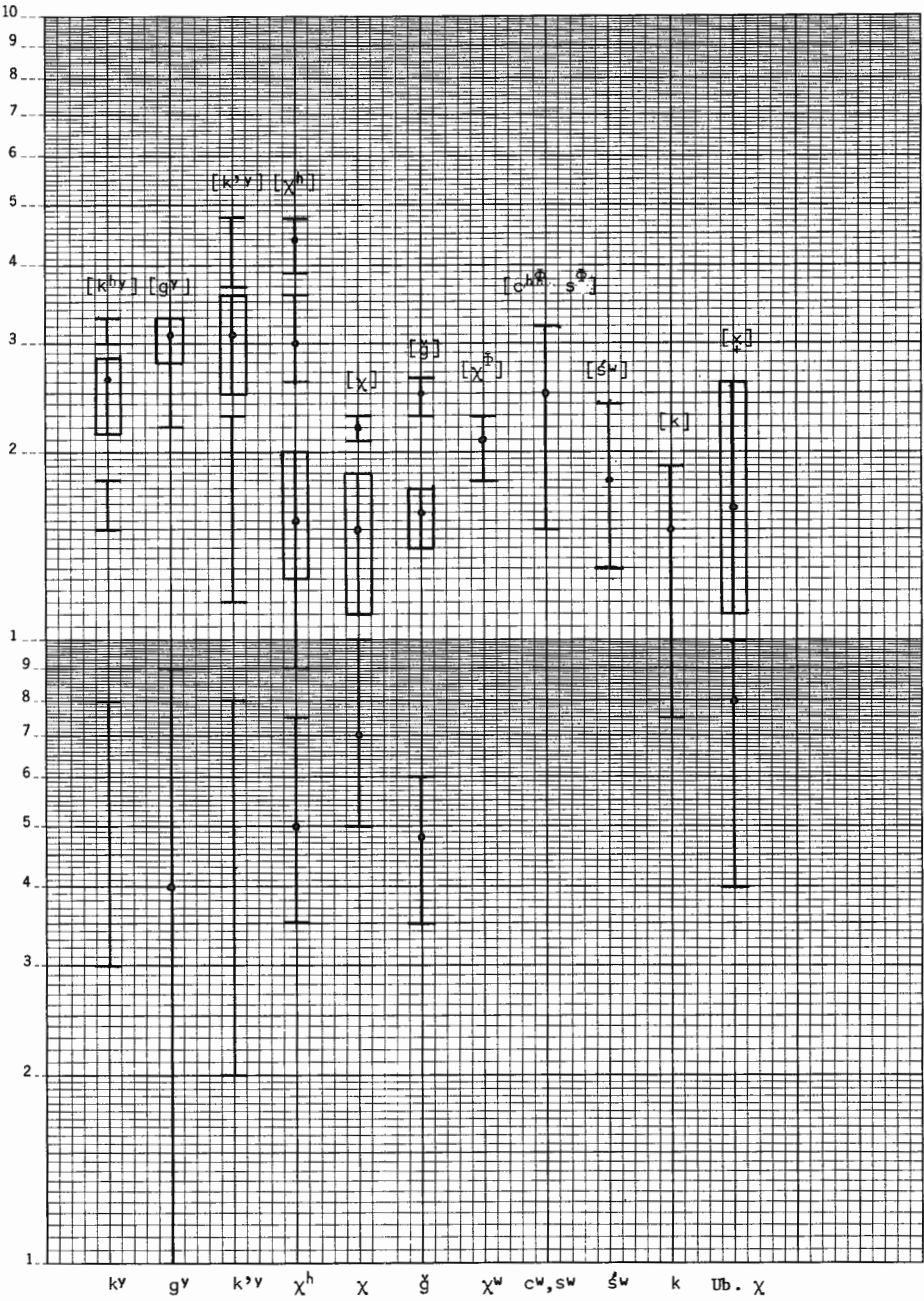
/ğ/ completely absent from Abkhaz and Abaza.

23.1. With this series I leave behind the coronals and begin the examination of segments made in the back of the oral cavity. For the most part these have acoustic ranges lower than those of the coronals and yet higher than those of the labials. Graph 5 starts on the left with Ubykh /k^y(=[k^{hy}]), g^y, k^y/ and Bzhedukh /χ^h, χ, ğ/. The acoustic energy for the stops is high, centering between 2.65 and 3.1 kHz, higher than for many of the coronals, but also showing low frequency energy as well, centering around the 0.4 kHz, something not seen for the coronals. Figure 15 shows Ubykh /k^y/ (+[k^{hy}]) (Leroy and Paris 1974:266-67), a clear palatal with dorsal articulation.

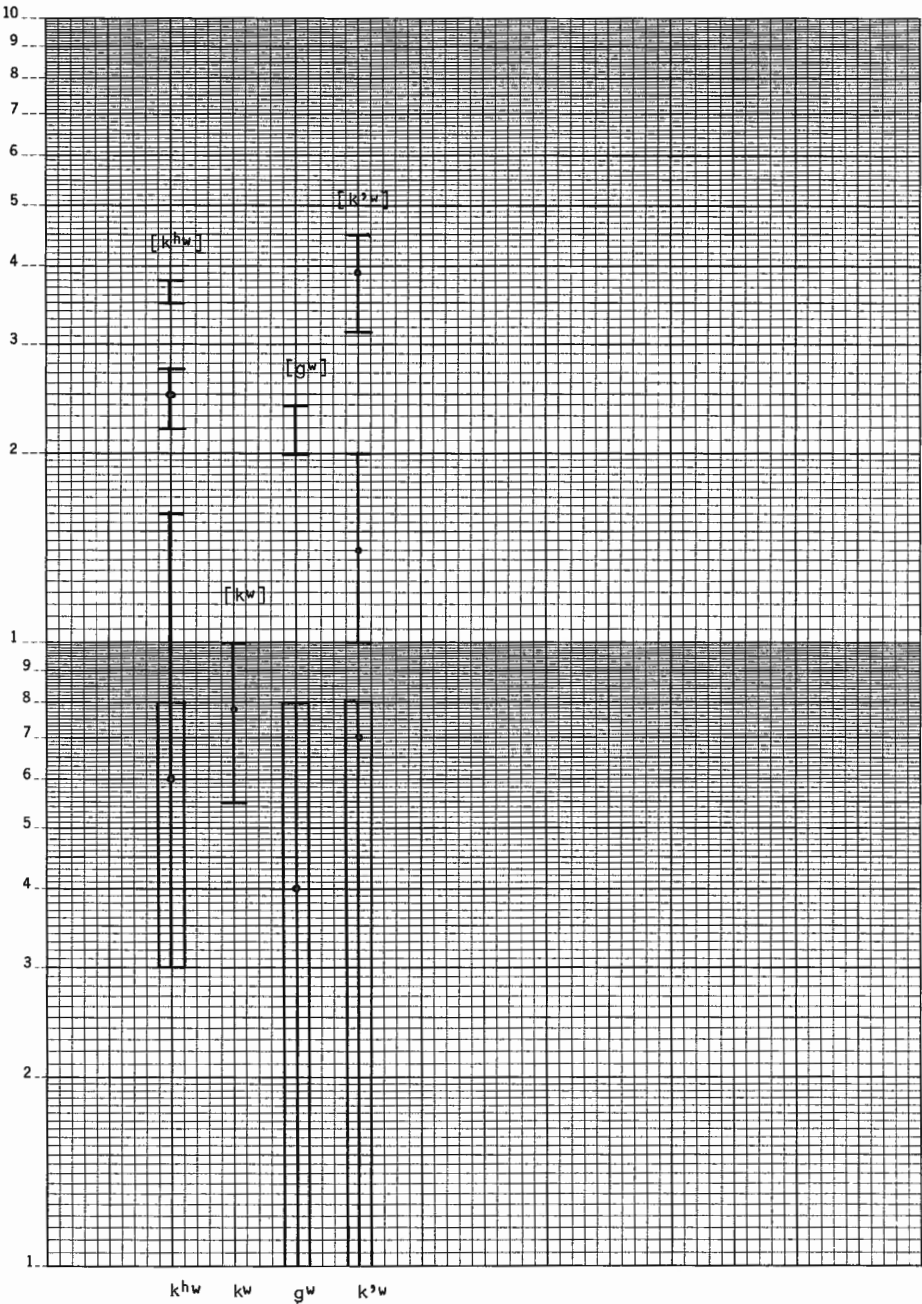
23.2. I have argued elsewhere (Colarusso 1983) that /χ^h/ is from an underlying /h/. Nevertheless in Shapsegh and Bzhedukh /χ^h/ is a stable acoustic form that differs from its unaspirated counterparts in having pseudo-formant structure that is more prominent at higher frequencies than are those in either /χ/ or /ğ/.

23.3. A striking peculiarity of the spirant usually noted as /χ/ in Ubykh is that it appears to be an advanced tongue root velar as figure 16 indicates (Leroy and Paris 1974:266). This may be a conservative state of affairs reflected more thoroughly in the “palatalized” (advanced tongue root) uvulars of Ubykh, Abkhaz, and Abaza (§27). The range of high frequency noise is wider for this segment than for Bzhedukh /χ/ graph 5. This is similar to the effect of [+ATR] upon Bzhedukh /ś^w/ (graph 4, column 6), as opposed to that of [-ATR] upon Ubykh /ś^w/ (graph 4, column 8).

23.4. One should note that the centers for the palatal and [+ATR] velar spirants both cluster about 1.60 kHz. Why these should be so alike, and why the palatal spirants should be so different from the palatal stop burst noise, are both unclear.



Graph 5 Palatals and Velars (Part 1)



Graph 5 Continued (Part 2)

Figure 15. Ubykh $k/y/$ ($=[k^{hy}]$)Figure 16. Ubykh $/\chi/$

24. Rounded or labialized palatal spirants and glide: $[\chi^{yw}, \chi^{\phi}, \text{q}]$ and $/y^w/$ [$+high, -back, (+labial)$ or $(+round), \pm consonantal, \pm continuant$]

$[\chi^{yw}] \leftarrow / \chi\text{-}w/$ in Kabardian.

$[\chi^{\phi}] \leftarrow / \chi^w/$ of velar spirants (?) in Ubykh and Abaza.

$[\text{q}] \leftarrow / \text{q}^w/$ in Abzhwi-Samurzkan Abkhaz.

$/y^w/ \leftarrow * / \text{q}^w/$ in Bzyb Abkhaz.

24.1. Kuipers (1960:26-27) has acutely noted the subtle contrast between a rounded velar spirant, $[\chi^w]$ in (4.i), and a palatal spirant that has undergone rounding due to phonetic process, $[\chi^{yw}]$ in (4.ii).

(4) Rounded velar spirant vs. rounded palatal spirant in Kabardian.

(i) $/\emptyset\text{-}\emptyset\text{-}\chi^w\text{-}wa\text{-}p\lambda\text{-}a/ \rightarrow [\chi^{w\partial}\text{-}p\lambda e]$

he-him-for-dynamic-look-in, 'he looks in for his sake'

(ii) $/\emptyset\text{-}\emptyset\text{-}\chi a\text{-}w\text{-}p\lambda\text{-}a/ \rightarrow [\chi^{yw\partial}\text{-}p\lambda e]$

he-it-in-dynamic-look-in, 'he looks inside it'

I have confirmed this contrast with my informant, but have been unable to subject the data to spectrographic analysis.

24.2. Figure 4 shows Ubykh $/\chi^w/$, which Leroy and Paris (1974) have taken for $/s^w/$ (cf. §§11.2-11.4). In articulation this sound is clearly dorsopalatal, as I have noted. In graph 5 it is contrasted with the acoustics of the other palatals at left and with Ubykh $/c^w, s^w/$ (both in one bar) and $/\acute{s}^w/$, both to its right. $/\chi^w/$ is acoustically distinct from all of these, particularly from instances of $/s^w/$, which has a much broader range of noise than the true palatal $/\chi^w/$. $/\chi^w/$ may perhaps be aligned with $/k^w, g^w, k'^w/$, rounded velars, for purposes purely on phonological symmetry, but there is no acoustic justification for doing so.

24.3. Abaza $/\chi^w/$ occurs in a few loans from Kabardian, such as $/\text{ʔ}^w a\acute{s}\eta a\text{-}m\acute{a}\chi^w a/$ 'hill-happy, blessed', i.e., 'Mt. Elbruz', Kabardian $/\text{ʔ}^w a\acute{s}\eta a\text{-}m\acute{a}\chi^w a/$ 'idem', beside native $/a\text{-}x^w \acute{a}\text{-}wa\text{ʔ}^w r\lambda/$ 'the-hill-blessed', 'idem', (Allen 1965:161). It is clear from tapes made by Allen¹⁵ of a

Tapanta Abaza speaker from Turkey that /χ^w/ is realized as /χ^φ/ (IPA [χ^φ]). The Abaza data were not of a quality suitable for spectrographic analysis. As with Ubykh, this sound might be assigned to the rounded velar series for purposes solely of phonological parallel, without phonetic justification.

24.4. Catford (1977a:190-91) suggests that Abkhaz /ś^w, ź^w/ are realized as [χ^φ, ʁ^β] ([χ^φ, ʁ^J] in his notation), respectively. This possibility has already been discussed in §16.

24.5. Catford (1972:680) attributes palatal rounding to Abaza /^{ʁw}/, thus /^{ʁʷ}/ . While this may be true for some speakers, from Allen's tapes of Tapanta I would prefer a form with back rounding, high or mid, [^{ʁw}] [^ʁ], respectively. Ashkharawa Abkhaz generally goes by the term 'Abaza' because it realizes its pharyngeals as pharyngeals, as does Abaza, so variations between [^{ʁʷ}], [^{ʁw}], and [^ʁ] may exist in it as well. In Abzhwi-Samurzakan, on the other hand, /^{ʁw}/ ranges from [u] to [^{ʁʷ}] (Catford 1972:680). In Bzyb, however, such processes of "emphatic" palatalization (Colarusso 1985:366; Trubetzkoy 1939:124) have taken an original */^{ʁw}/ over fully to a /y^w/. When this new glide loses its rounding in certain variants, as in (5), then a /y/ is the result and not some pharyngeal. In some cases this /y/ undergoes the glide-vowel metathesis characteristic of nonround glides in this language (Colarusso 1979:313). Forms are from Bgažba (1964:112-13).

(5) Bzyb Abkhaz */^{ʁw}/ → /y^w/ → /y/

(i) /a-ləmħa-r-y^wə/ → /a-ləmħa-r-əy/
the-ear-by-companion, 'ear-ring'

(ii) /a-y^wən-ča-rà/ → /a-yn-ča-rà/
the-house-drive out-infinitive, 'to drive some one out of his home'

(iii) /à-ž^wy^wan/ → /à-ž^wyan/
the-heaven, sky

(iv) /a-č^wy^wàn/ → /a-č^wyàn/
the-stake, picket

(v) /a-t^wəy^wa/ → /a-t^wəya/
the-horn, antler

(vi) /y^wa-ž^w-y-ž^wa-ba/ → /ya-ž^w-y-ž^wa-ba/ (Kaladakhwar subdialect)
2-10-and-10-thing, '30 things'

This rephonemicization to a nonpharyngeal accounts for Catford's comment that none of his Bzyb informants showed any trace of pharyngeal articulation for this segment. Unfortunately data were not available for spectrographic analysis.

25. Plain velar stops: /k, g, k'/

[+high, +back]

In Abkhaz and Abaza, where they are [k^h, g, k'].

/k/ (= [k]) and /k'/ in Bzhedukh West Circassian clusters.

25.1. The plain velar stops are confined to the Abkhaz-Abaza branch, the southern portion of the family, where they are ubiquitous. Dumézil (1964:185) claimed rare instances of all three for Ubykh, but in (1965:228, 230) this claim had weakened to just /k, g/ and in (1975:12, §11.5) only /k'/ is noted for a few borrowings. Ubykh sides with Circassian in this feature.

25.2. Outside of Abkhaz-Abaza only the Bzhedukh dialect of West Circassian shows phonemically voiceless unaspirated /k/ and glottal ejective /k'/ in clusters: /šk-/ or /šχ-/ 'to eat', /š'k'a/ 'calf (up to one year old)' (Kuipers 1975:56, 53; 1963:77). Graph 5 exhibits the acoustics of several icons of Bzhedukh /k/. Its stop burst range is from 0.75 to 1.90 kHz, centering about 1.5 kHz. This is in keeping with the range for English unrounded back /k/ (before /a/) as shown in Cooper et al. (1952:600).

26. Rounded velar stops, spirant, and glide: /k^{hw}, k^w, g^w, k'^w, χ^w, w/

[+high, +back, +round]

/k^w/ only in Shapsegh and Bzhedukh West Circassian.

All others ubiquitous.

26.1. The rounded glide /w/ and its variants have been treated in §3, 4. Graph 5 is completed with stop burst plots for Bzhedukh and Ubykh rounded velar stops. These display remarkable diversity, but share at least low energy noise below 1.0 kHz. This is in keeping with the stop burst range for English /k/ before /ɔ, o, u/, which is slightly rounded (Cooper et al. 1952:600). No velar /χ^w/ was available for analysis, but presumably it would differ markedly from Ubykh palatal /χ^w/.

26.2. Figure 17 shows Ubykh /k^w/ (Leroy and Paris 1974:267).

**Figure 17. Ubykh /k^w/**

26.3. Hewitt (1979:257) reports labialization for this series in Abzhwi-Samurzakan, rather than rounding. This labialization of the velars is in marked contrast with the lip-rounding of the uvulars. Unfortunately, no data were available for analysis.

27. Palatalized uvulars: /qʸ, qʸʸ, xʸ, ɣʸ/

[-high, +back, +ATR]

/qʸ/ in Ubykh only.

/qʸʸ, xʸ, ɣʸ/ in Ubykh, Abkhaz, and Abaza.

27.1. These sounds are typologically extremely unusual (Colarusso 1979:309, 310, 318). In the feature system of Chomsky and Halle (1968) one would expect palatalization to be [+high] and for a palatalized uvular, therefore, to be automatically a simple velar. These palatalized uvulars, however, are indeed uvulars. The tongue is bunched up and so forms an occlusion that extends up onto the anterior portion of the velum while narrowing the anterior oral cavity (Colarusso 1979:309; 1975:219-91; 1974; Catford 1977a:290). Figure 18 (Leroy and Paris 1974:268) shows this remarkable effect.



Figure 18. Ubykh /xʸ/

27.2. The chief acoustic result of advancing the tongue root (Graph 6) in this articulation is to produce noise ranging from 1.45 to 4.0 kHz, with centers ranging from 2.30 to 2.85 kHz. Perkell (1972) has suggested that expanding the pharynx by advancing the tongue root results in an intensified and lowered first formant in a vowel, and goes on to suggest that similar effects might be observed in consonants. Only /ɣʸ/ shows any effect of this sort, although the high frequency noise in the other members of this series may be thought to recapitulate the high frequency amplification already seen in Bzhedukh /ś/ and Ubykh /χ/. Lowering and enhancement of the first formant in adjacent vowels do, however, give these vowels a palatal color and directly account for the palatal quality of these consonants.

27.3. Worthy of note is the fact that these segments are opposed to plain and pharyngealized uvulars as well. Thus, one must have both [+ATR] and [+constricted] pharynx (Klatt and Stevens 1969; Perkell 1971) as antagonistic features and one cannot merely assume that a feature [expanded (pharynx)] is sufficient for mere two way contrasts as Lindau (1979) did. Lindau's "universal" that "no language contrasts more than two [degrees of pharyngeal width]" (1979:176) is, as is the case with so many putative universals, simply wrong, being based on languages in which pharyngeal activity is very limited. That the two features are antagonistic can be seen from the Bzyb Abkhaz and Abaza forms in (6) which contrast with the Abzhwi forms by dissimilating between the two features. In (6.i) [+CP] (for "constricted pharynx" henceforth) is produced to enhance the opposition to [+ATR], and in (6.iii, iv), [+ATR] results to enhance the opposition to [+CP] (Colarusso 1979:309-310).

(6) Forms illustrating dissimilations involving [+ATR] and [+CP].

- (i) P(roto)-A(bkhaz)-A(baza) */-xʲγʷà/ → */-xʲγʷà/ → Bzyb Abkhaz /-xʲxʷà/ 'to regret'
- (ii) P-A-A */-xʲγʷà/ → */-xʲγà-/ → Abzhwi Abkhaz /-xʲʕa-(ga-)/ → [xʲaaga] 'idem'
- (iii) P-A-A */-q̣qʷà/ → Bzyb /-x̣qʷà/ 'cover', cf. /a-x̣čʷəx̣qʷà/ 'pillow case' (probably /a-x̣-čʷə-x̣qʷà/ the-head-support(?) -cover)
- (iv) P-A-A */-q̣qʷà/ → Abaza /-q̣qʷà/ 'idem', (also /q̣čʷq̣qʷà/ 'pillow case')
- (v) P-A-A */-q̣qʷà/ → Abzhwi /-xʕʷa/ → [xq̣a] 'cover'

27.4. As strange as these segments are they are not unique to these languages. Palatalized uvulars contrast with plain uvulars, plain velars and palatalized velars (palatals?) on the phonetic level in Nanay Tunguz (Menges 1969:179).

28. Uvular stops and spirants: /qʰ, q, qʷ, x, xʰ, γ/

[-high, +back]

/q/ as a phonemic contrast only in Shapsegh, Bzehdukh and Hakuchi West Circassian.

/qʷ/ absent in all West Circassian save for Hakuchi (Kuipers 1963).

/x/ only in Ubykh dialect (Dumézil 1965:266-69)

28.1. Graph 6 shows the acoustics of this series. The range of the stop bursts and spirant pseudo-formants for these sounds is surprisingly close to that for the plain velar graph 5, the chief difference between them being that the uvulars have their energy concentrated at a lower range, so that their acoustic centers are lower than in the velars. Figure 19 shows Bzyb /x/ (Bgažba 1964:92). Figure 20 shows both Ubykh /x/ and /xʷ/



(Leroy and Paris 1974:267), the latter being the rounded counterpart of the plain uvular.

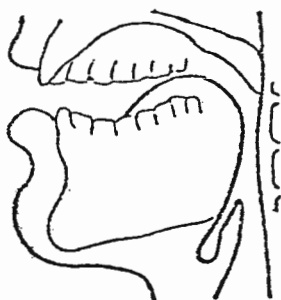


Figure 19. Bzyb Abkhaz /x/



Figure 20. Ubykh /x/ and /x^w/

29. Rounded uvular stops and spirants: /q^{hw}, q^w, q^w, x^w, x^w, γ^w/

[-high, +back, +round]

/q^w/ as a phonemic contrast only in Shapsegh, Bzhedukh, and Hakuchi West Circassian.

/q^w/ absent in all West Circassian, save for Hakuchi (Kuipers 1963). /x^w/ only in Ubykh dialect (Dumézil 1965:266-69).

29.1. Graph 6 shows the acoustics of this series. Lip rounding has resulted in a lower range for stop bursts and pseudo-formants than that for the plain uvulars. High frequency noise, which plays a minor role in the plain uvulars, has been entirely damped out in the rounded ones. Figure 20 shows that the point of articulation for this series is exactly the same as that for the plain uvulars.

29.2. Jakobson et al. (1951:31-36) posited a feature [flat] which included rounded, retroflexed, and pharyngealized segments, all of which shared an acoustic tendency toward lower targets than those found in [-flat] counterparts. This acoustic lowering is an accurate characteristic of the effects of rounding. Pharyngealization and retroflexion, on the other hand, are acoustically more complex. This is particularly evident with pharyngealization (or some variety of it), as may be seen in the following two sections.

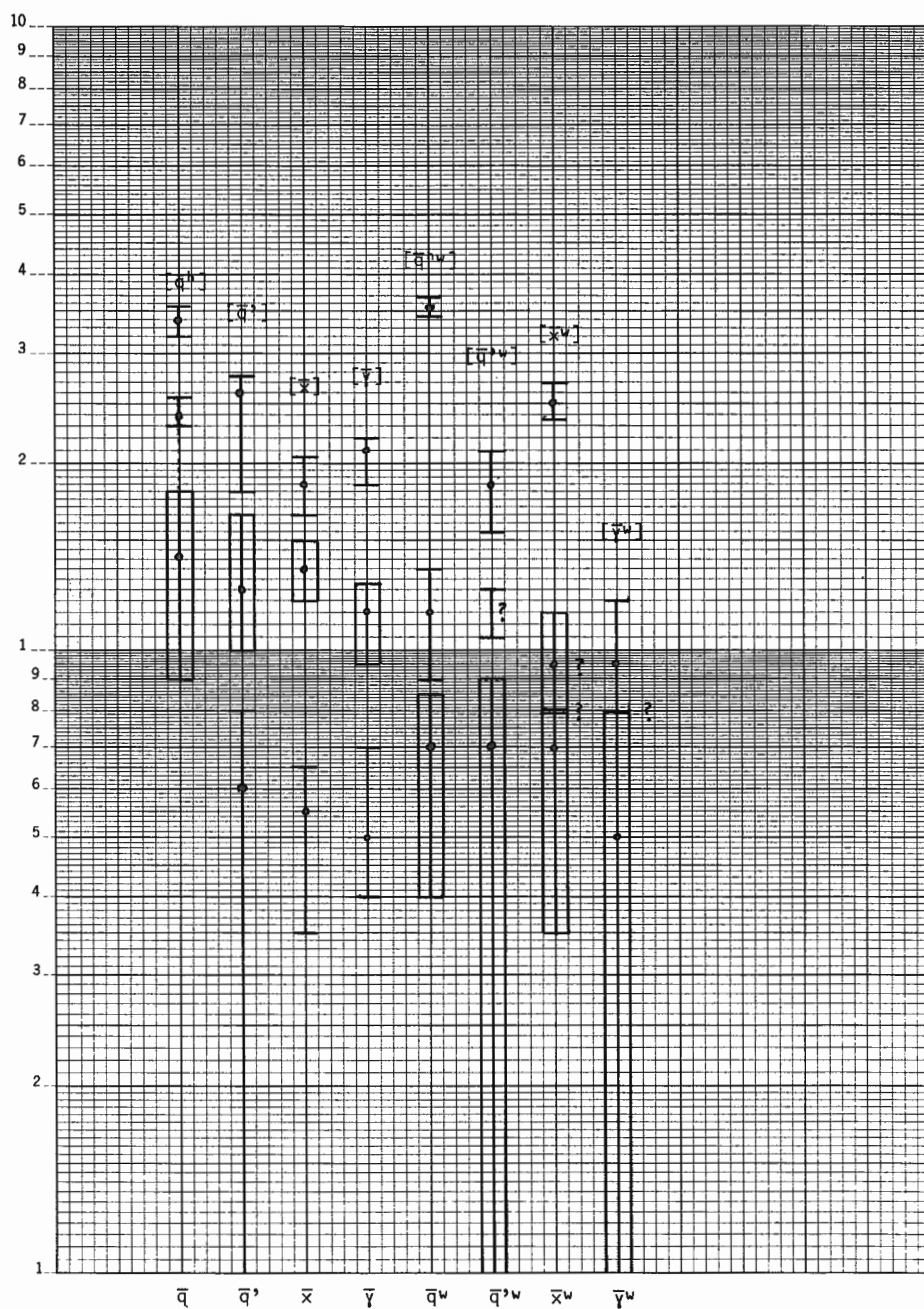
30. Pharyngealized uvulars stops and spirants: /q̣, q̣', x̣, γ̣/

[-high, +back, +CP]

Whole series in Ubykh.

/q̣/ chiefly in Abaza.

/q̣/ or /x̣/ in variation in Bzyb and Ashkharawa Abkhaz (Bgažba 1964; Lomtadze 1954).



Graph 7 Pharyngealized Uvulars, Plain and Unrounded

30.1. Graph 7 exhibits the complex acoustics of this remarkable series. Intense noise may be observed in the range from 0.90 to 1.80 kHz, with the co-occurrence of both high and low noise beyond this range. Only /q̃/ ([q̃^h]) lacks the low energy noise. With their strongest centers about 1.15 to 1.40 kHz these segments exhibit emphatic palatalization (cf. §24.5) (Colarusso 1985:366, for native American language examples of this effect). Clearly the acoustics of this series is distinct from that of the rounding of its predecessor. The feature [flat] must be discarded, (Colarusso 1979:310; 1975:220-21; 1974).

30.2. Figure 21 shows Bzyb Abkhaz /x̃/ (Bgažba 1964:90); figure 22 shows Tapanta Abaza /q̃/ (Bgažba 1964:92); figure 23 shows both Ubykh /x̃/ and /x̃^w/ (Leroy and Paris 1974:268), which differ only by rounding (§31). In fact, in every language with pharyngealized uvulars each of the uvulars has a rounded counterpart (§31).

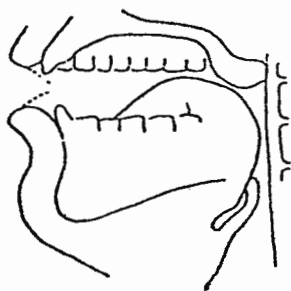


Figure 21. Bzyb Abkhaz /x̃/



Figure 22. Tapanta Abaza /q̃/



Figure 23. Ubykh /x̃/ and /x̃^w/

30.3. Comparison of figures 21-23 with the plain uvulars in figures 19 and 20 shows clearly that the former set have retracted tongue roots and

consequent narrowing of the pharynx. The exact nature of this pharyngealization will be taken up in §32. The Ubykh segment has secondary retroflexion, but on a universal basis these two gestures must be distinguished (Bhat 1974).

30.4. A point of interest is that Abaza has two pharyngealized uvular stops /q̤/ and /q̤ʷ/ (§31) that are realized as voiceless aspirated uvulars with moderate pharyngealization. Bgažba's x-ray tracing conclusively proves this point. In fact Abaza represents one end of a spectrum that extends from Tapanta, through Ashkharawa Abkhaz into Bzyb, the latter two showing variation between stop and spirant realizations of these two segments, with Bzyb tending to favor spirants. As far as I can determine, all who have worked upon Abaza other than Bgažba have missed this pharyngealization. For example, Allen (1955:129, table 1, fn. 3) notes that the series /q, qʷ/ is defective in /qʲ/ and thus is contrasted with the complete series /qʲ, q', q'ʷ/. In fact this series is /q̤, q̤ʷ/ and it has no defect, since */q̤ʲ/ would involve the mutually antagonistic feature specifications [+CP] and [+ATR].

30.5. Pharyngealized uvulars are found in the Northwest Caucasian family in Lakk (Trubetzkoy 1931:34), Kubachi (Deeters 1963:32; Magometov 1963:47-50), dialects of Dargwa (Magometov 1963:47-50), Tabasaran (Magometov 1965:66-71), Ginukh (Hinukh) (Lomtadze 1963:41-48), Dido (Imnajšvili 1963:16, 19-20), and in Archi (Kibrik et al. 1977, vol. 1:224, 333-48). Outside the Caucasus such sounds are found only in the Interior Salishan languages of North America, such as Coeur d'Alene (Colarusso 1985:367; Kinkade 1967:233).

31. Pharyngealized uvular stops and spirants, rounded: /q̤ʷ, q̤ʷʰ, x̤ʷ, ɣ̤ʷ/

[-high, +back, +CP, +round]

Whole series in Ubykh.

/q̤ʷ/ chiefly in Abaza.

/q̤ʷ/ and /x̤ʷ/ in variation in Bzyb and Ashkharawa and Abkhaz.

31.1. This series dramatically illustrates the inadequacy of the old feature [flat] (Colarusso 1979:310). Graph 7 shows these complex sounds. Rounding has resulted in a shift of the most prominent portion of their spectra from the zone running between 0.90 and 1.80 kHz to that running between 0.10 and 0.90 kHz. Higher frequency stop burst and spirant noise has been damped out in this series. Their prominent low frequency band accounts for their remarkably dark and turgid quality. They are distinguished from plain rounded uvulars (§29) by having greater low frequency noise and by the presence of some higher frequency pseudo-formants. Figure 23 serves double duty for both Ubykh /x̤ʷ/ and /q̤ʷ/. Figure 24 shows Bzyb /x̤ʷ/ (Bgažba 1964:90), which is clearly a uvular spirant with retracted tongue root, much as may be seen in

figure 21. Noteworthy in figure 24, however, is the suggestion that [+round] is realized with lip approximation rather than simple rounding. No one who has worked on this particular form of Abkhaz has commented on this effect, however, but simply referred to these sounds as though they had the same sort of rounding as nonpharyngealized rounded uvulars.

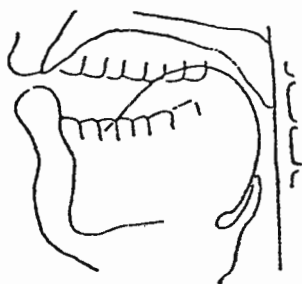


Figure 24. Bzyb Abkhaz / \bar{x}^w /

31.2. While plain pharyngealized uvulars are widespread in the Northeast Caucasian family, rounded ones are not, being found only in Archi (Kibrik et al. 1977). In Interior Salishan Coeur d'Alene, however, rounded pharyngealized spirants occur as counterparts for every non-rounded uvular (Colarusso 1985; Kinkade 1967).

32. Pharyngeals, plain: / h , ʕ / [+CP]

Absent from Ubykh.

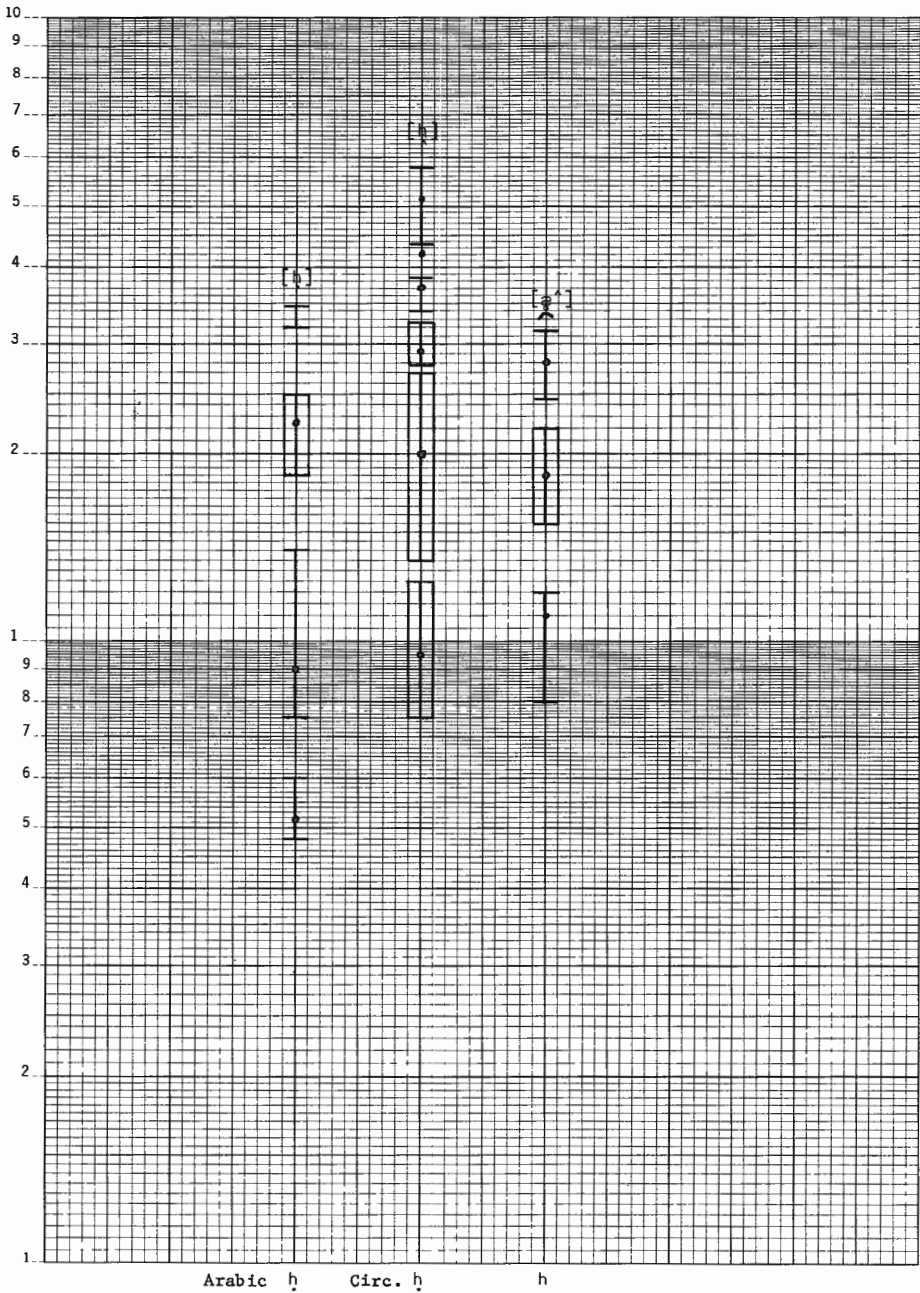
/ ʕ / only in Arabic borrowings into Circassian.

32.1. Graph 8 shows the acoustics of Bzhedukh / h /. These are contrasted with a true voiceless pharyngeal spirant to the right taken from Lebanese Arabic (Klatt and Stevens 1969:212). Clearly the formant structure of the Circassian sound is different from that of a true pharyngeal. The Circassian sound is also louder. The acoustic analysis corroborates my impression that the Circassian "pharyngeals" are more raucous and turbid sounding than true pharyngeals. Such raucous "pharyngeals" have also been noted for the Wakashan language Nootka (Colarusso 1985:367; Sapir in Jacobsen 1969:125-26). Catford (1977a:163) contrasts West Circassian ("Adyghe") / h / as an ordinary pharyngeal ("faucal") with a deeper / ʕ /, but such is not the case for Bzhedukh, both segments being lower than true pharyngeals, as graph 8 proves at least for / h /. Catford correctly emphasizes, however, the distinction between these "ventricular" sounds (as he calls them) and true faucal or retracted tongue root pharyngeals of the Arabic type. S. M. Kodzasov of Moscow University has pointed out to me (personal

communication) that the Northwest Caucasian “pharyngeals” (which he terms “emphaticals”) can be made with the tongue stuck out, so that retraction of the tongue root is clearly not at play.

32.2. Pending x-ray studies I can only conjecture that this series is the consonantal counterpart to what Chomsky and Halle (1968:314-15) posit as a feature [covered], found in some African vowel-harmony systems. The steps involved in the development and production of these sounds, both historically and to some extent in the phonological derivations at the synchronic level, seem to be as follows (Colarusso 1985): The feature value [+low] appears to involve an opening of the oral cavity, manifest in such articulations as those of the classic low vowels. This opening of the oral tract is mutually antagonistic to the radical oral constriction underlying the feature value [+consonantal] (Chomsky and Halle 1968:302). [+low] can occur, therefore, only in [-consonantal] sounds that are not made in the oral cavity. This restriction leaves only the larynx, which is “minus everything” with regard to point of articulation and perhaps the pharynx, which might be argued to fall outside the zone of the oral cavity proper, as non-vocalic candidates for [+low]. I have argued elsewhere (Colarusso 1981:502-552; 1979:312-14; 1975:384-408) that laryngeals can be associated with [+low]. If this is true of [+CP] segments as well, then such a [+CP, +low] sound would involve not only retraction of the tongue root, but lowering of the root down into the pharynx as well. Such lowering would bring the radix down against the epiglottis, and this would bring the latter into approximation with the “adytus” or opening into the larynx. Such a low pharyngeal would have a significant amount of “adytal” friction. The retracted tongue root forms seen in §§30-31 suggest such articulations. Once established in a language, such multiply articulated pharyngeals might lose the tongue root retraction and retain only the adytal friction as their realization. Such “adytals” seem to be used to realize the pharyngeals in Circassian and Abaza and the voiceless pharyngeals in Abkhaz. The voiced member has a particularly “bleating” quality (Catford 1977a:163).

32.3. Such adytals are realized in the larynx and are subject to further modification to pure laryngeals. This is precisely the case with /ʕ/ in Abkhaz (Colarusso 1979:313) where the general tendency is to take underlying /ʕ/ over to [f] with spirant quality or even to [a] with regular voicing (cf. Dumézil 1975:14, §12). As [a] in syllable coda this segment causes syllable nuclei to assimilate, taking on [+low], so that the result is a long syllable nucleus, [aː]. In syllable onset [a] persists (Dumézil 1967:10, §2).



Graph 8 Pharyngeal, “Adytal,” Laryngeal

33. Pharyngeals, rounded: /ħ^w, ʕ^w/

[+CP, +round]

Only in Abkhaz-Abaza.

33.1. Rounded variants of the adytals or pharyngeals are found in Abkhaz and Abaza. I have already discussed the palatal realization of /ʕ^w/ in §24. No data was available for analysis.

34. Laryngeals, stop and spirant: /ʔ, h/

/ʔ/ = [-labial, -coronal, -anterior, -high -back, -continuant]

/h/ = [-labial, -coronal, -anterior, -high, -back, +continuant, +low]

Only in Circassian, save for Hakuchi (Kuipers 1975:5), and in Circassian loans into Abaza (with a few native exceptions) and Ashkharawa Abkhaz.

/h/ in Ubykh.

34.1. /h/ has a laryngeal realization in two Ubykh forms, /hày/ 'no' and /həndà/ 'today, now' (Vogt 1963:129-30). In Kabardian (Kuipers 1960:21-22) /h/ shifts to /ɦ/ when the preceding segment is voiced. In Circassian this segment is restricted largely to the plural morpheme /-ha/ suffixed to nouns and verbs. In West Circassian this morpheme is realized as [χɛ] or [χ], save for Shapsegh and Bzhedukh dialects where a [χ^h] emerges (see §23). In this instance of the plural morpheme /h/ does not undergo metathesis with the following /a/ and therefore behaves as if it were [+consonantal] (Kuipers 1960:23, 34). In the third person plural /-y-ha-/ 3rd. p. plural 'them, they, their', the segment acts as though it were a glide, i.e., [-consonantal], and undergoes metathesis with its following vowel, the result being [yɑ·] or [yɑ·] (cf. Abkhaz /ʕ/ in §32.3). The resulting long low vowel is stable, not reducing when unstressed, and in this regard is contrasted with a long vowel resulting from /aa/ (Colarusso 1983). Other instances of [ɑ·] that persist when unstressed are also plausibly construed as being /ha/ with /h/ [-consonantal]. In initial position, as in /a-/ 'the', the realization is usually [ɑ·], but can also be [hæ] or [ha], both in West Circassian (my own observations) as well as in Kabardian (Jakovlev 1948:343). This suggests that such morphemes might better be represented as /ha-/ or that there is a rule that fills an empty syllable onset before a nucleus [σ_aσ] with an /h/. In any event the phonology of /h/ is much more interesting than its phonetics.

34.2. Graph 8 shows Ubykh /h/. It is realized as a spirant onset to the formants of the following vowel as is to be expected. Comparison with the adytal /h/ shows the latter is much closer to a laryngeal as realized before a low vowel (the form used was Ubykh /hày/ 'no') than it is to a conventional pharyngeal.

34.3. The glottal stop acts throughout as though it were [+consonantal], never undergoing metathesis. In fact both /ʔ/ and /ʔ^w/

can be shown to have come from Proto-Circassian */q'/ and */q'w/, respectively and might still be best ranked with the uvulars (Kuipers 1975:4; Balkarov 1970:258).

34.4. A few Abkhaz dialects that have been subjected to strong Turkish influence realize their /ħ/ as [h] (cf., Dumézil 1967:9, §1, fn. 3, for one of these).

35. Laryngeals, rounded: /ʔ^w, h^w/

[+round], minus all else save for [+continuant] for /h^w/ and [+consonantal]

/ʔ^w/ in all Circassian save for Hakuchi West Circassian.

/h^w/ in Besleney East Circassian as spoken in Turkey (Paris 1974b:121).

35.1. These segments have simple lip-rounding. No metathesis takes place between either laryngeal and any following vowel, so both are [+consonantal]. In Besleney /h^w/ is the reflex of the usual East Circassian /χ^w/, and in fact may merely be a surface realization of the same.

35.2. /h^w/ was not available for analysis and /ʔ^w/ merely exhibits the formants of the following rounded vowel with transient glottalic trill or abrupt formant onset. The transient glottalic trill with mid vowel formant targets is the realization in such forms as /χ'-əy-t'ʔ^w/ 'man-plural-connective-two', 'two men'. I have not plotted the form since the results would have been the same as plotting the formants for various round vowels.

36. Laryngeals, palatalized: /ʔ^y/

[+high, +consonantal], minus all else.

Only in Abadzakh West Circassian.

36.1. This odd phoneme is the reflex of earlier */č'y/ or */k'y/ (Balkarov 1970:137, 189) for some speakers of Abadzakh West Circassian, others having the original or a spirantized version, /š'y/ (cf. §17). For example 'child' can be /č'yāaλa/, /š'yāaλa/ or /ʔ'yāaλa/, the last being realized as [ʔəyλε]. No forms were available for analysis, but my field work leads me to expect a glottal trill or abrupt onset to a fronted vowel formant structure as a likely acoustic pattern.

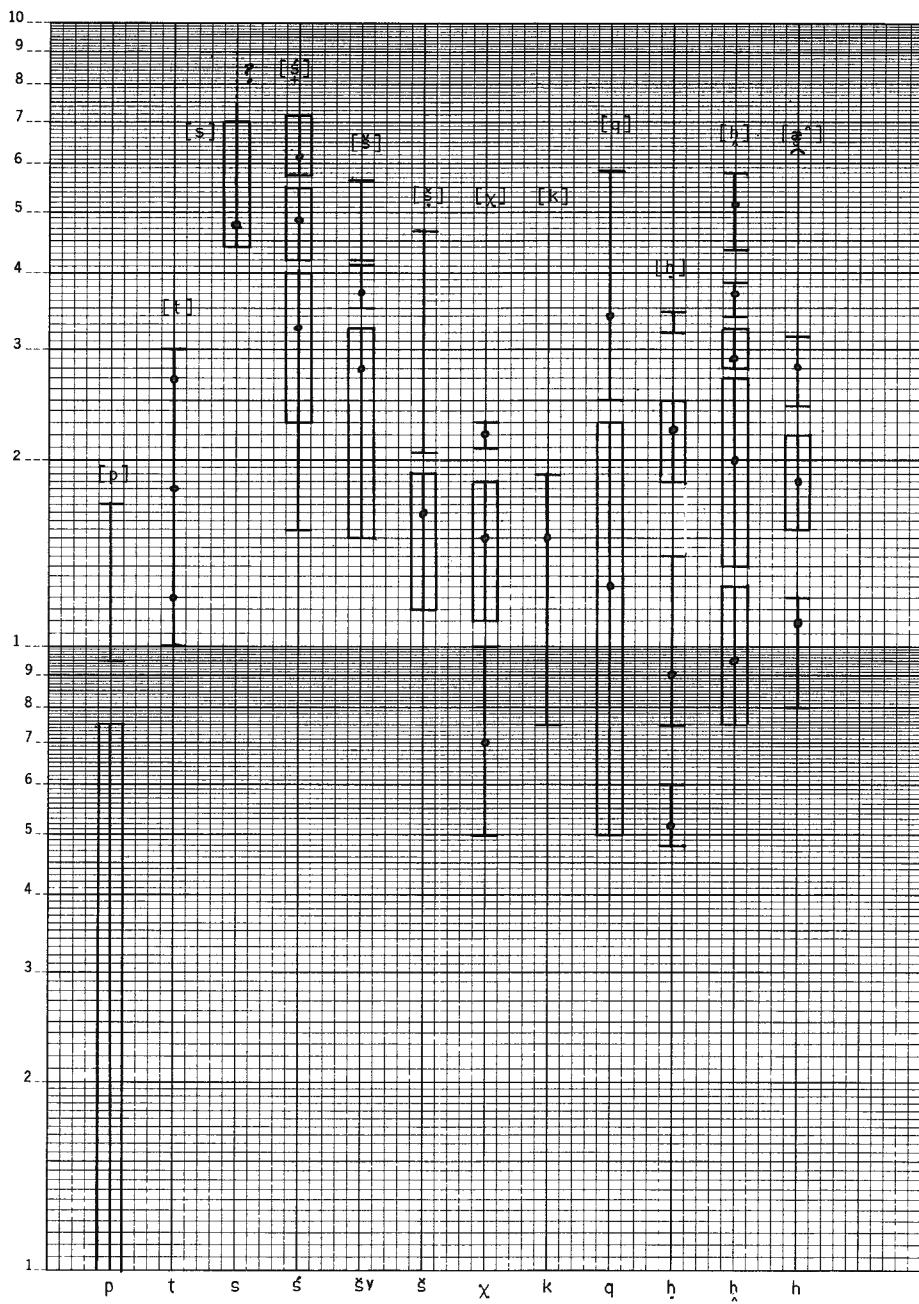
37. Conclusion: Phonetic Space.

37.1. Psychological spaces are used for two purposes by psychologists. One is to plot the subjective interplay of semi-independent or at least distinct dimensions of perception. This is the case with the well-known color spindle which shows the interrelationships between the distinct perceptual parameters of hue, intensity and saturation. The other use is to capture invariant relationships between distinct parameters that can themselves undergo varying realizations. The musical torus is an example

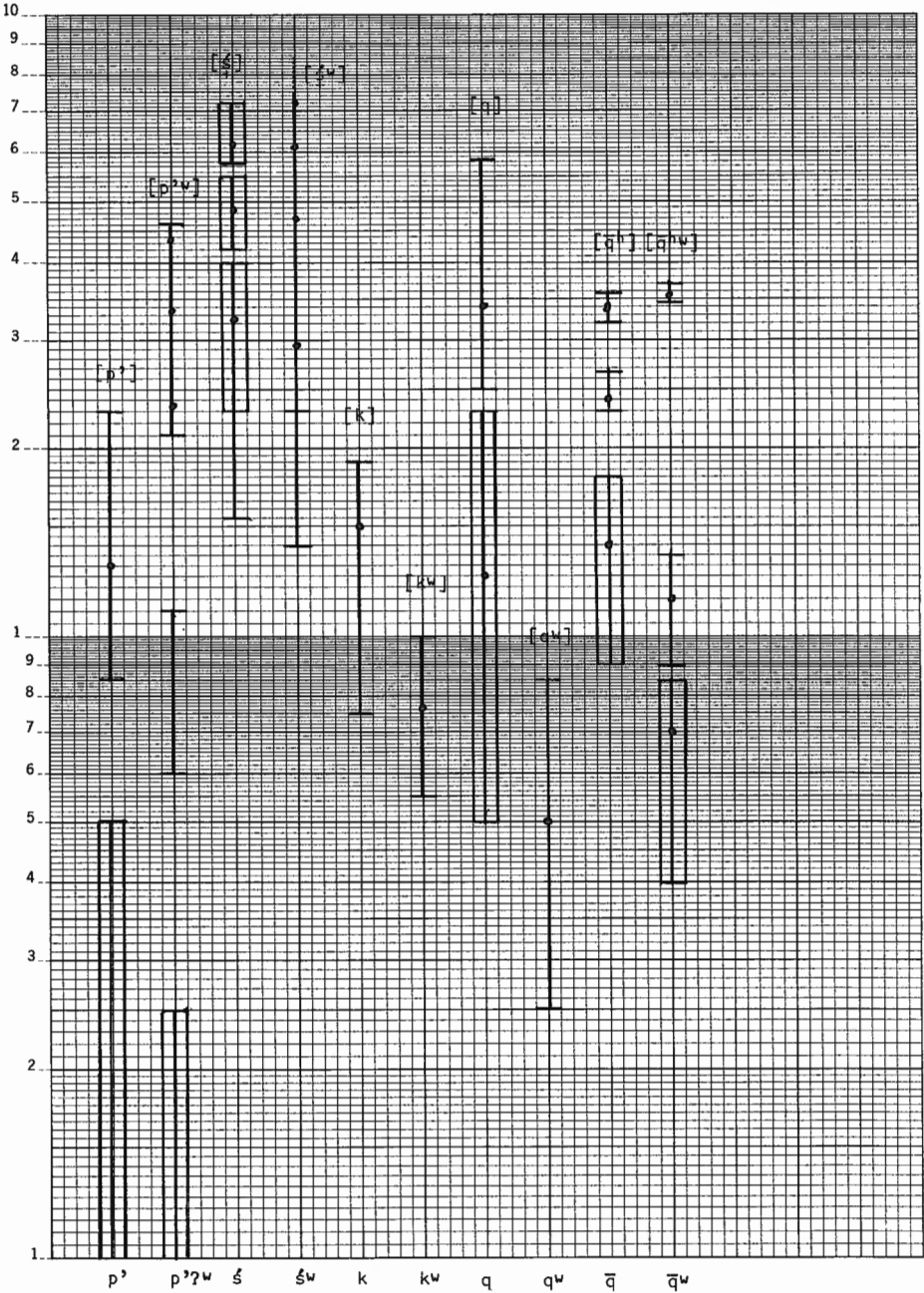
of this type of space in which major keys are linked with corresponding minor keys in a way constant and independent of the note chosen for the tonic, i.e., the starting point for the major scale in the chosen key. I propose here that all the preceding data can be summarized in a complex multidimensional phonetic space of predominantly the first type, though aspects of the second might also prove useful for certain pairs of phonetic parameters. As a multidimensional space of more than three dimensions phonetic space can not be envisioned easily, though graphs 9-12 can be seen as two dimensional sub-sections through the complete hyperspace, as it is technically called.

37.2. Graph 9 shows a partitioning of the frequency parameter as a function of points of articulation. Graph 10 shows the interaction of some of these frequency zones with the secondary modification of rounding.¹⁶ Graph 11 is like 10 except that the modification is that of pharyngealization. Graph 12 completes the series with the modification of advanced tongue root. Graph 9 is quite thorough, thanks to the richness of the Northwest Caucasian languages. The other graphs must be viewed as preliminary efforts at tracing the relationships between various phonetic parameters. Other relationships could and must be plotted, even utilizing the present data, but this is work for a future study (Colarusso, in preparation). The present roughly delimited hyperspace might be viewed as the first effort to give exact features and structure to the abstract perceptual space proposed by Marcel A. A. Tatham (1984) to account for the perceptual capabilities exhibited by speakers when perceiving acoustic input.

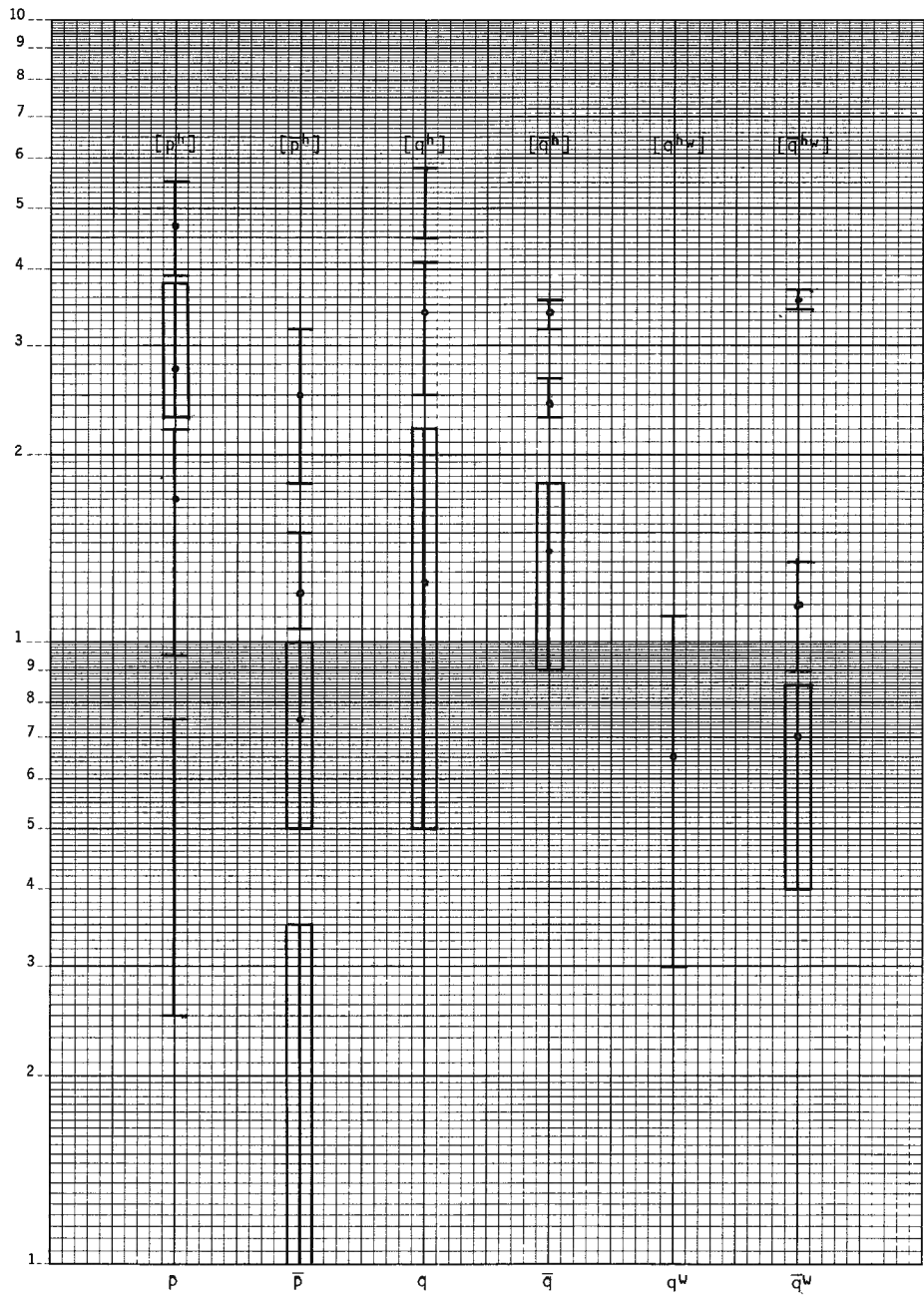
37.3. Should a quantal theory of speech perception be desired, and the present data is mixed on this score, then one might partition this proposed hyperspace into stable zones of equivalence, much as Tatham himself suggests. To do this one might have to resort to folding the hyperspace into such invariant zones and determining membership in them by a judicious choice of parameters. These parameters would then be the dimensions of the hyperspace containing the folded phonetic hyperspace. Such enclosing dimensions are called co-dimensions and a natural way to perform the folding is via the mathematical theory called catastrophe theory (Thom 1983; 1975; Saunders 1980).¹⁷



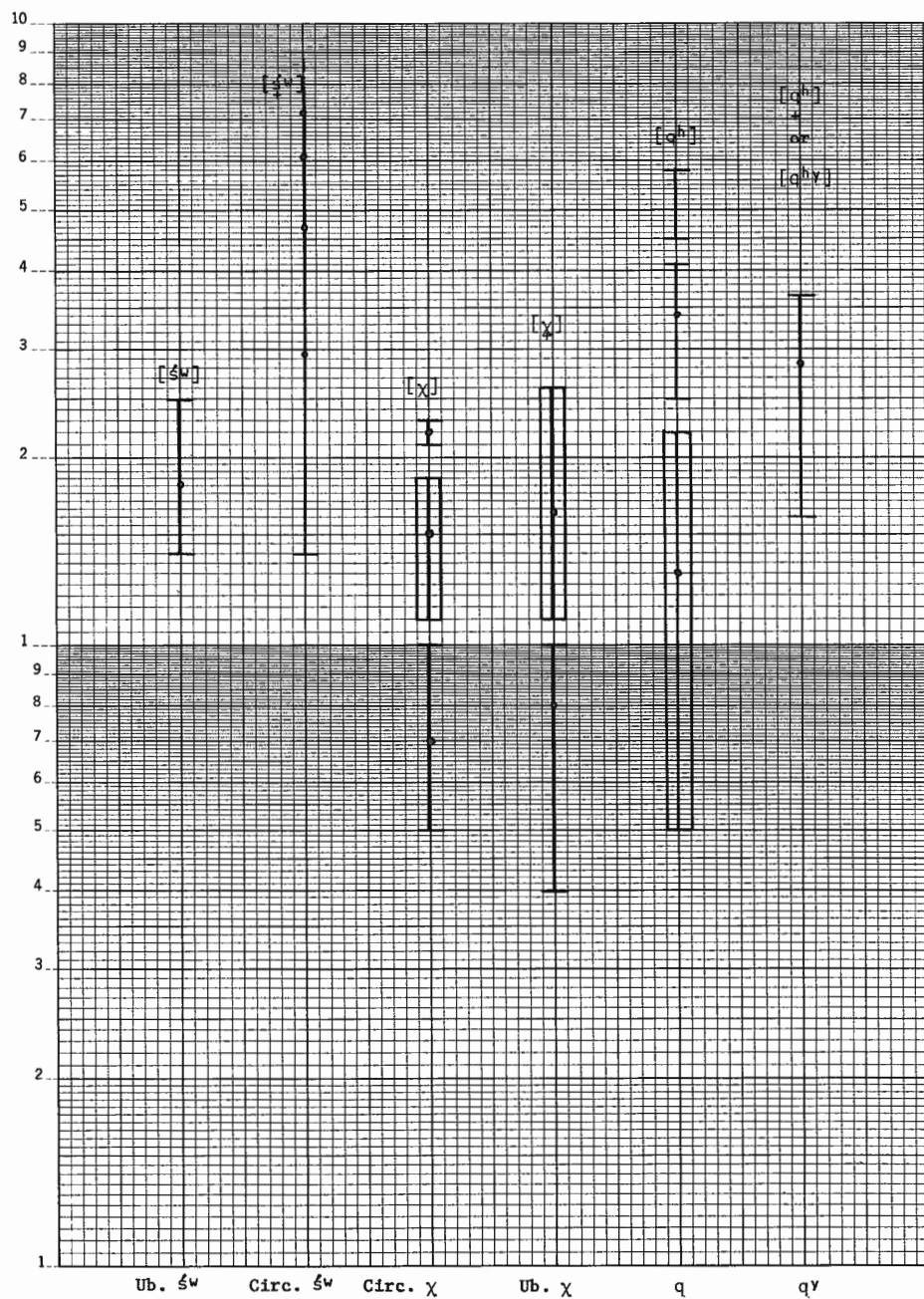
Graph 9 Point of Articulation Partitioning of Frequency



Graph 10 The Effects of Rounding



Graph 11 The Effects of Pharyngealization



Graph 12 The Effects of [+ATR]

Notes

¹The historical process at work in the production of such odd systems appears to have been one in which vocalic features have migrated from the syllable nucleus out to the syllable margins where they became consonantal features, thus Proto-Northwest Caucasian */t'q'o/ → West Circassian /t'ʔwə/, East Circassian /t'ʔwə/, Ubykh /t'q'wə/ (with secondary /a/), Bzyb Abkhaz /y'wə/, Abaza /ɕ'wə/ both from Proto-Abkhaz-Abaza */t'ɕwə/, all 'two'.

²For example, the Chadic Language Margi has eighty-seven consonants (Colarusso 1975:viii-ix) and some of the Koisian languages have just as many, such as ꞎHòǎ/ (/t'←k^hǎǎ/) also with eighty-seven (Gruber 1972).

³One can readily show that the rounded, palatalized, pharyngealized, etc., sounds are unit phonemes (Colarusso 1983:12-15; Catford 1972:682).

⁴Hans Vogt generously supplied the tapes and elicitation lists for Ubykh, not to mention years of advice and support in my study of this language.

The spectrograms were made at MIT, where Keith North allowed me access to the sonagraph exclusively for my own purposes for days on end.

The graphs are logarithmic, since pitch is perceived logarithmically. In other words, the ratio between frequencies determines perceived pitch, not the difference between them.

⁵Dumézil's corrections to Vogt's dictionary (Dumézil 1965:197-269; 1964) must be viewed with great circumspection. For example, he views /aa/ as leading to a stable [a], but in fact, Ubykh /aa/ is not stable. It can be realized as [æ], [a] or [ɑ], just as with Circassian /aa/ or /ah/, or Abkhaz /aɕ/. The critical cue for "long a" (which is not significantly longer than the other vowels (Catford, personal communication)) is syllable coda and only secondarily length: CaaCV → Ca'.CV or Ca.CV, vs. CaCV → CaC.CV. Thus Dumézil often posits [q] when /a/ is called for merely because an icon of a word has a particularly low vowel realization. Vogt (personal communication) informed me that /aa/ was unstable in Esenç's speech, perhaps an effect due both to the moribund state of the language and to Circassian influence whereby unstressed /aa/ → /a/. I have listened carefully to Vogt's tapes and in nearly every instance I side with Vogt against Dumézil. Clearly, however, some corrections are warranted or plausible. Vogt, for example, overlooked /ya-/ 'to spy upon', and /caqapə/ may be a variant of what is clearly /caqapə/ 'dried corn' on Vogt's tapes.

⁶Catford's derivation of this item is more complex: T. /kalpak/ → Ub. */qal~paq/ → /qal~ba/ → /qaba/ → /baq'ǎ/. Catford himself, however, suggested another Turkish loan wherein initial /ka-/ has dropped, just as I have assumed in my derivation of 'hat.' His derivation is: T. /kabuğa/ → (Ub. */baɣ'wa/ →) /baɣ'a- 'to pick, clean fruit', with irregular loss of rounding my interpretation.

⁷Ub. /pla-/ 'to look at' must be a Circassian borrowing. For Ub. /f/ ← Proto-Northwest Caucasian */p^h/ cf. Ub. /faɕ'ǎ/ 'nose, beak', West Circassian /p^ha/ 'nose, tip, front'; Ub. /-fa-/ preverb 'at the tip of something', West Circassian /-p^ha-/ 'idem'; Ub. /-fa-/ preverb 'downwards' as in /-fa-t'^w-/ 'down-climb, move' West Circassian /-p^hǎ-(λaλa-)/ 'down-(dangle)'; Ub. /fa-/ 'to tear out (weeds, plants, etc.)', West Circassian (Bzhedukh) /-p^hǎ-ɕ^{hy}ǎ-/ 'sever-tear'. Thus the Ubykh develop-

ment was PNWC */p^hə-λ-a-/ 'point-look-at' (for */-λ-/ cf. West Circassian /-λaaɣ^w-/ 'to see') → */fλa-/ → */fλa-/ → */fa/ → /va-/, the last shift because no */f/ exists in Ubykh.

⁸I am indebted to Micheal Job (personal communication) for suggesting this analysis of Shapsegh.

⁹In the poorly attested dialect of Ubykh the labial component has dominated so that /t^w, d^w, t'^w/ → /p, b, p'/.

¹⁰The tracings of Leroy and Paris have come under some criticism for, I think, three reasons: (1) they neglect detail; (2) it is not clear that they represent stills coinciding with the main portion of the articulation; (3) they produce strange results (cf. §11). They are the only tracings of Ubykh and so I use them, but also they (1) represent all the salient features; (2) with the exception of "s^w" they seem to be synchronized with the main articulation; (3) they contain no confusion, but rather reveal a hitherto neglected phoneme (cf. §§11, 22).

¹¹'Ten,' listed as /z^w/ in Leroy and Paris, is a misprint for /z^wə/ (Vogt 1963:222, entry 2353).

¹²The Lo-Kuban dialect (Genko 1955:29, 48) is the only language in the entire family to contrast two full series of rounded coronals; /č^w, ž^w, č'^w, š^w, ž^w/ ~ /č^{yw}, ž^{yw}, č'^{yw}, š^{yw}, ž^{yw}/. All other Abaza dialects have only one series which represents a collapse of three original ones: */t^w, d^w, t'^w/, */č^w, ž^w, č'^w, š^w, ž^w/ and */s^w, ž^w/.

¹³In many Soviet works this series is unmarked and the retroflexed or velarized series is marked by a following 'ə'.

¹⁴Noteworthy is the Apsuy subdialect of Ashkharawa Abkhaz which has only one rounded coronal series that can be realized as alveopalatal, laminal palatoalveolar or apical palatoalveolar affricates and spirants, all within the speech of the same individual (Lomtatz 1954:4-5; contra my earliest report (Colarusso 1979:311)).

¹⁵W. S. Allen gave this tape to Calvert Watkins, who in turn passed it on to me.

¹⁶/p^{rw}/ is taken from Jakobsen et al. (1951:50, fig. 6). There was no frequency scale for this form so I have estimated one based upon the formant targets of the following vowel, which is [o], as the form is clearly Bzhedukh West Circassian /p^{rw}a/ = [p^{rw}o] 'shell'.

¹⁷See Wildgen (1982) for an application of catastrophe theory to semantic analysis.

References

- Abitov, M.L., et al. 1957. *Grammatika kabardino-čerkeskogo literaturnogo jazyka*. Moscow.
- Allen, W. S. 1956. "Structure and System in the Abaza Verbal Complex." *Transactions of the Philological Society* 1956:127-76.
- _____. 1965. "An Abaza Text." *Bedi kartlisa, revue de kartvelologie* 19-20:159-72.
- Baddeley, John F. 1908. *The Russian Conquest of the Caucasus*. (Reissued 1969 by Russell & Russell, New York.)
- Bagov, P. M. 1969. *Kubano-zelenčukskie govory kabardino-čerkeskogo jazyka*. Nal'chik.

- Balkarov, B. X. 1970. *Fonetika adygskix jazykov*. Nal'chik.
- Bgažba, X. S. 1964. *Bzybskij dialekt abxazskogo jazyka*. Tblisi.
- Bhat, D. N. S. 1974. "Retroflexion and Retraction." *Journal of Phonetics* 2:233-37.
- Bouda, Karl. 1939. "Tscherkessische volkslieder aus Armavir." *Zeitschrift der deutschen morgenländischen Gesellschaft* 93:339-62.
- Catford, J. C. 1942. "The Kabardian Language." *Le Maître phonétique* 78:15-18.
- _____. 1972. "Labialization in Caucasian Languages with Special Reference to Abkhaz." In *Proceedings of the Seventh International Congress of Phonetic Sciences*, eds. A. Rigault and R. Charbonneau, pp. 679-82. The Hague.
- _____. 1977a. *Fundamental Problems in Phonetics*. Bloomington.
- _____. 1977b. "Mountain of Tongues: the Languages of the Caucasus." *Annual Review of Anthropology* 6:283-314
- Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York.
- Colarusso, John. 1974. "Consonants with Advanced Tongue Root in the Northwest Caucasian Languages." In *NELS V, Papers from the Fifth Annual Meeting, North Eastern Linguistic Society*, eds. Ellen Kaisse and Jorge Hankamer, pp. 153-61. Harvard University, Linguistics Department.
- _____. 1975. "The Northwest Caucasian Languages: a Phonological Survey." Unpublished Ph.D. dissertation. Department of Linguistics, Harvard University.
- _____. 1979. "Phonemic Contrasts and Distinctive Features: Caucasian Examples." In *The Elements: a Parasession on Linguistic Units and Levels, Including Papers from the Conference on Non-Slavic Languages of the USSR*, eds. Paul R. Clyne et al., pp. 307-321. Chicago: Linguistic Society.
- _____. 1981. "Typological Parallels between Proto-Indo-European and the Northwest Caucasian Languages." In *Amsterdam Studies in the Theory and History of Linguistic Science IV Current Issues in Linguistic Theory*, vol. 16, part I, *Bono Homini Donum: Essays in Historical Linguistics, in Memory of J. Alexander Kerns*, eds. Yoël L. Arbeitman and Allan R. Bomhard, pp. 475-557.
- _____. 1982. "Western Circassian Vocalism." *Folia Slavic* 5:89-114.
- _____. 1983. "Compensatory Lengthening, Abstractness and the Logic of Synchronic Phonology: an Examination of Bzhedukh West Circassian." Unpublished MS.
- _____. 1985. "Pharyngeals and Pharyngealization is Salishan and Wakashan." *International Journal of American Linguistics* 51:366-68.
- _____. 1992. "How Many Consonants does Ubykh Have?" In *Caucasian Perspectives*, ed. B. George Hewitt, pp. 145-156. Munich: Lincom Europa.
- _____. (in preparation). "A First Mapping of 'Phonetic Space.'"
- Cooper, Franklin S. et al. 1952. "Some Experiments on the Perception of Synthetic Speech Sounds." *The Journal of the Acoustical Society of America* 24:597-606.
- Deeters, Gerhard. 1963. "Die kaukasischen Sprachen." In *Armenisch und kaukasischen Sprachen*, eds. G. Deeters, G. R. Solta and V. Inglisian, *Handbuch der Orientalistik* 1, *Der Nahe und der Mittlere Osten*, 7, ed. B. Spuler, pp. 1-79. Leiden.
- Dumézil, Georges. 1964. Review of "Hans Vogt. — *Dictionnaire de la langue oubykh*." *Bulletin de la Société de Linguistique de Paris* 59:189-92.

- _____. 1965. *Documents anatoliens sur les langues et les traditions du Caucase*, III, *Nouvelles études oubykhs*. Paris.
- _____. 1967. *Documents anatoliens sur les langues et les traditions du Caucase*, V, *Études Abkhaz*. Paris.
- _____. 1970. "Notes sur le parler des Abazaz d'Anatolie." *Folia Linguistica* 2:275-78.
- _____. 1974. "Notes d'étymologie et de vocabulaire sur le Caucasic du nord-ouest, 8-10." *Journal asiatique* 262:19-29.
- _____. 1975. *Le verbe oubykh*. Paris.
- Gamkrelidze, T. V. 1966. "A Typology of Common Kartvelian." *Language* 42:69-83.
- Gamqrelize T. and Givi Mačavariani. 1965. *Sonant'ta sist'ema da ablaut'i kartvelur enebši*. [The sonant system and ablaut in the Kartvelian languages] Tbilisi.
- Genko, A. N. 1955. *Abazinskij jazyk*. Moscow.
- Gruber, J. S. 1972. "≡Hōā Vocabulary." Unpublished elicitation list for specimen reel.
- Halle, Morris and G. N. Clements. 1983. *Problem Book in Phonology*. Cambridge, MA.
- Halle, Morris and Kenneth N. Stevens. 1969. "On the Features 'Advanced Tongue Root'" *MIT Quarterly Progress Report* 94:209-215.
- Henderson, E. J. A. 1970. "Acoustic Features of Certain Consonants and Consonant Clusters in Kabardian." *Bulletin of the School of Oriental and African Studies, University of London* 33:92-106.
- Hewitt, B. George. 1979. *Lingua Descriptive Studies, Abkhaz*. Amsterdam.
- Imnajtšvili, D. S. 1963. *Didojskij jazyk, v sravnenii s ginuxskim i xvaršijskim jazykami*. Tbilisi.
- Jacobsen, W. H., Jr. 1969. "Origin of the Nootka Pharyngeals." *International Journal of American Linguistics* 35:125-53.
- Jakobson, Roman, C. Gunnar, M. Fant, and Morris Halle. 1951. *Preliminaries to Speech Analysis*. Cambridge, MA.
- Jakovlev, N. F. 1930. "Kurze Übersicht über die Tscherkessischen (Adygeischen) Dialekte und Sprachen." *Caucasica* 6:1-19.
- _____. 1948. *Grammatika literaturnogo kabardinočerkesskogo jazyka*. Moscow.
- Jongman, Alfred, Sheila E. Blumstein and Aditi Lahiri. 1985. "Acoustic Properties for Dental and Alveolar Stop Consonants: a Cross-Language Study." *Journal of Phonetics* 13:235-51.
- Keraševa, Z. I. 1957. *Osobennosti šapsugskogo dialekta adygejskogo jazyka*. Maykop.
- Kibrič, A. E., S. V. Kodzasov, I. P. Olovjannikova, and D. S. Samedov. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka, I, leksika, fonetika*. Moscow.
- Kinkade, M. Dale. 1967. "Uvular-pharyngeal Resonants in Interior Salish." *International Journal of American Linguistics* 33:228-34.
- Klatt, Dennis H. and Kenneth N. Stevens. 1969. "Pharyngeal Consonants." *MIT Quarterly Progress Report* 93:207-216.
- Kuaševa, T. X. 1969. "Terskie govory." In *Očerki kabardino-čerkesskoj dialektologii*, ed. M. A. Kumaxov. Nač'čik.
- Kuipers, Aert H. 1960. *Phoneme and Morpheme in Kabardian*. The Hague.

- _____. 1963. "Proto-Circassian Phonology." *Studia Caucasica* 1:56-92.
- _____. 1975. *A Dictionary of Proto-Circassian Roots*. Lisse.
- Kumaxov, M. A. 1961. "Fonetičeskie osobennosti malkinskogo govora kabardino-čerkeksskogo jazyka. In *Voprosy izučeniya iberijsko-kavkazskij jazykov*, ed. È. A. Bokarev. Moscow.
- _____. 1969. "Kubanskij dialekt." In *Očerki kabardino-čerkeksskoj dialektologii*, ed. M. A. Kumaxov. Nač'ik.
- Ladefoged, Peter. 1964. *A Phonetic Study of West African Languages*. Cambridge University Press.
- Laycock, D. C. 1965. *The Ndu Language Family (Sepik District, New Guinea)*. Canberra.
- Lehmann, W. P. 1955. *Proto-Indo-European Phonology*. Austin, TX.
- Leroy, Christine and Catharine Paris. 1974. "Étude articulatoire de quelques sons de l'oubrykh d'après film aux rayons X." *Bulletin de la Société Linguistique de Paris* 69:255-86.
- Lindau, M. 1979. "The Feature 'Expanded'." *Journal of Phonetics* 7:163-76.
- Lomtadze, E. A. 1963. *Ginuxskij dialekt didojsskogo jazyka*. Tbilisi.
- Lomtadze, K. 1954. *Ašxaruli dialekt'i da misi adgili sxva apxazur-abazur dialekt'ta šoris* [the Ashkharawa dialect and its place among the other Abkhaz-Abaza dialects]. Tbilisi.
- Magometov, A. A. 1963. *Kubačinskij jazyk*. Tbilisi.
- _____. 1965. *Tabasaranskij jazyk*. Tbilisi.
- Mamrešev, K. T. 1969. "Baksanskij dialekt." In *Očerki kabardino-čerkeksskoj dialektologii*, ed. M. A. Kumaxov. Nač'ik.
- Menges, Karl H. 1968. "Die tungusischen Sprachen." In *Tungusologie*, eds. W. Fuchs et al., *Handbuch der Orientalistik*, ed. B. Spuler, erste Abteilung, V. Band, 3 Abschnitt, pp. 21-256. Leiden.
- Oden, Gregg C. 1978. "Integration of Place and Voicing Information in the Identification of Synthetic Stop Consonants." *Journal of Phonetics* 6:83-93.
- Paris, Catherine. 1974a. *Textes chapsough d'Anatolie (tcherkess occidental), avec vocabulaire aux texts*. Paris.
- _____. 1974b. *Système phonologique et phénomènes phonétiques dans le parler besney de Zennun köyü (tcherkess oriental)*. Paris.
- Perkell, Joseph. 1971. "Physiology of Speech Production: a Preliminary Study of Two Revisions of the Features Specifying Vowels." *MIT Quarterly Progress Report* 102:123-37.
- Pike, Eunice V. 1964. "The Phonology of the New Guinea Highlands Languages." In *New Guinea, the Central Highlands*, ed. J. B. Watson, special publication *American Anthropologist* 66.4, part 2:121-32.
- Pike, Kenneth L. 1967. "Tongue Root Position in Practical Phonetics." *Phonetica* 17:129-40.
- Rogava, G. V. and Z. I. Keraševa. 1966. *Grammatika adygejskogo jazyka*. Krasnodar and Maykop.

- Santerre, Laurent and Ching Y. Suen. 1981. "Why Look for a Single Feature to Distinguish Stop Consonants?" *Journal of Phonetics* 9:163-74.
- Saunders, P. T. 1980. *An Introduction to Catastrophe Theory*. Cambridge, MA.
- Šagirov, A. K. 1961. "Fonetičeskie osobennosti malkinskogo govora kabardino-čerkeskogo jazyka." In *Voprosy izučenija iberijsko-kavkazskix jazykov*, ed. É. A. Bokarev. Moscow.
- _____. 1969. "Malkinskij govor." In *Očerki kabardino-čerkeskoj dialektologii*, ed. M. A. Kumaxov. Nal'čik.
- Stevens, Kenneth N. 1972. "The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data." In *Human Communication, a Unified View*, eds. E. E. David, Jr., and P. B. Denes. New York.
- Stewart, J. M. 1967. "Tongue Root Position in Akan Vowel Harmony." *Phonetica* 16:185-204.
- Stevens, Peter. 1960. "Spectra of Fricative Noise in Human Speech." *Language and Speech* 3:32-49.
- Tarnóczy, Thomas. 1948. "Resonance Data Concerning Nasals, Laterals and Trills." *Word* 4:71-77.
- Tatham, Marcel A. A. 1984. "Towards a Cognitive Phonetics." *Journal of Phonetics* 12:37-47.
- Thom, René. 1975. *Structural Stability and Morphogenesis*. Reading, MA.
- _____. 1983. *Mathematical Models of Morphogenesis*. New York.
- Trubetzkoy, Nikolai. 1931. "Die Konsonantsysteme der ostkaukasischen Sprachen." *Caucasica* 8:1-52.
- _____. 1939. *Grundzüge der Phonologie*. Göttingen.
- Vogt, Hans. 1960. Unpublished elicitation list for Ubykh tapes.
- _____. 1963. *Dictionnaire de la langue oubykh*. Oslo.
- Wildgen, Wolfgang. 1982. *Catastrophe Theoretic Semantics*. Amsterdam.

Assertive Verb Forms in Lak⁰

Victor A. Friedman

In my article on the typological significance of status in the Lak verbal system (Friedman 1984), I accepted the terminologies of Žirkov (1955) and Burčuladze (1979) for the purpose of examining Lak status categories in the broader context of Balkan and Caucasian languages. Implicit in my acceptance of that terminology was the acceptance of certain assumptions about grammatical meanings associated with those terms unless otherwise specified. Thus, for example, I used the term *confirmative* (based on the Russian *utverditel'nyj* 'affirmative' and *podverditel'nyj* 'confirmative') for the synthetic finite verbal forms based on the participle in *-ss[a]-*, and I cited some examples of typical usage without being able to enter into greater detail, due to the paucity of information in the available sources. I also attempted to draw some limited parallels with forms called *confirmative* in Balkan languages, e.g., Bulgarian, Macedonian, and Turkish. In this article dedicated to Academician Akači Šanidze, I wish to examine these Lak screeves in greater detail, and I will try to characterize their basic meaning, which, as I hope to justify here, I will label *assertive*.

I will be primarily concerned with the synthetic, finite, aspectually unmarked, nonmodal present and past screeves of the type *čičajssar*, *-a*, *-u* ('write', present), *čivčussar*, *-a*, *-u* ('write', past), and to a lesser extent *čičajssija*, *-v* ('write', imperfect). In addition to the two Russian terms mentioned above, these forms have been described by the Georgian term *mṭṭṭicebiti* 'affirmative, assertive'.¹ The available sources on Lak, however, do not go much further than the label in defining these forms. Burčuladze (1979:244) compares them with the English construction in *do* as in the following example:

- (1) *Na čičajssar čağar*
I do write a letter.

Xajdakov (1961:116) observes that the present assertive is the standard screeve used in proverbs and cites a number of examples, including the following:

- (2) *Kä"valin lavgunni kunu laččul čimus naču x~unu qqabučajssar.*
Having gone to the Kaaba, garlic does not return sweet.

Murkelinskij (1971) generally renders the affect of the Lak assertive into Russian by using the adverb *dejstvitel'no* 'really', although he also uses *imenno* 'precisely' and *opredelenno* 'definitely', as in the following examples:²

- (3) *Na čivčussar. =Ja napisal (imenno).* (Murkelinskij 1971:180)
I wrote (precisely) [it].
(4) *Ina čivčussar. =Ty (dejstvitel'no) napisal.* (Murkelinskij 1971:180)
You (really) wrote [it].
(5) *Na zunčissara. =Ja (opredelenno) budu rabotat'.* (Murkelinskij 1971:181)
I (definitely) will work.

Beyond these, there are no explicit observations on the assertive other than, e.g., Xajdakov's that the plain and assertive imperfect are "*ves'ma blizki po značeniju drug drugu*" (Xajdakov 1953:12, cited in Burčuladze 1979:195) 'extremely close to one another in meaning.'³

The chief peculiarity of the present and past assertive is that they are used in what appear to be two distinct contexts and functions. On the one hand, these assertives function in colloquial speech as a type of emphatic, i.e., the speaker is purposefully emphasizing the truth of the proposition, as in example (6), where the surrounding context makes it clear that the speaker is greatly excited and is trying to impress his listeners with the urgency of the fact that the time to act has indeed arrived:

- (6) *Nanu! Čunssar!* (Murqčilinskij 1980:49)
Come on! It is time!

On the other hand, and this is not observed in the grammars, the present assertive functions in expository prose as the most common present tense screeve. Example (7) is typical in this respect:

- (7) *Agar maq qiva žuralij čičajssa bux~urča, slovar'danuvu qivagu žura kkakkan buvnu bussar.* (Xajdakov 1962:17)
If a word has two written forms, both forms are shown in the dictionary.

Consider also in this light the contrast between examples (8) and (9) as well as the statement made in example (10):

- (8) *A'kinssakssa mašinarttu baqqašivrijn buvnu, cila čumal qus tajla qqadurkssar.* (Murqqilinskij 1981:117)
 Due to the absence of the necessary cars, the goods were not sent in time.
- (9) *Χaqinu čanssa mašinarttu bija. Qus tajla dukkan žušča qqax~urna.* (Murqqilinskij 1981:117)
 There were too few cars today. We didn't send the goods.
- (10) *Iširal stil'danuvu asar kičlagan čülü buvssa maqru, kalimarttugu išla qqadajssar.* (Murqqilinskij 1981:119)
 In the business (practical) style, emotive and decorative words and phrases are not used.

These three examples are taken from the section on style in the seventh/eighth-grade grammar used in Lak schools (Murqqilinskij 1981). The context of the contrast between examples (8) and (9) is Murqqilinskij's explanation that the style of formal writing is different from the style of relaxed speech. He gives (8) as an example of how we would explain a situation in an official report, and he gives (9) as an example of how we would describe the same situation when talking with our friends. As can be seen, the official style in (8) uses a past assertive like the present assertive in the expository sentence in (7), whereas the colloquial version in (9) uses an unmarked imperfect and a simple preterite. Example (10) is another expository statement like (7) and (8). It is significant because it states a principle that applies to example (8) (and also to [7] and [10] itself), viz., the avoidance of highly colored language in business and expository, i.e., formal and neutral, style. This appears to establish a curious contrast: on the one hand the assertive can be said to be marked for something like 'personal emphasis' in colloquial speech, but on the other hand this same assertive is used in formal speech and writing as the neutral tense form. Before attempting to explain these facts, I shall adduce one more example:

- (11) *Graždan dā"vilul čumal Uxssavnil čuluxmur Kavkazulal zunttavu ukunssa iš x~ussar.* (Murqqilinskij 1981:22)
 During the civil war in the mountains of the North Caucasus such an incident occurred.

The form *x~ussar* 'occurred' is a past assertive used in the opening sentence of a short story. The story continues to use past assertive forms to set the scene, and then switches to a simple past form *uvkuna* 'he said' to introduce the first dialogue. After this the narration continues in the simple past. Example (11) is neither colloquial nor official but is taken from belles-lettres. The purpose here could be to invoke the neutrality of

expository prose, e.g., as if a news item or historical event were being reported in a formal context. Or the purpose could be to impress the reader, i.e., as if the author were saying: "This *really* happened..." This second possibility seems less likely since the assertive continues to be used for the entire first paragraph of the story until the introduction of the dialogue.

How, then, can the meanings of 'personal emphasis' and 'neutrality' be reconciled? One possible explanation is suggested by the Turkish enclitic *-dir*, etc., which functions both as a type of emphatic and as a copulative particle. The principle functions of *-dir* can be illustrated by the following examples from Lewis (1967:97):

(12a) *Vesika kasadadır.*

Writing/Formal Speech: The document is in the safe.

Informal speech: The document is surely in the safe/The document *is* in the safe.

(12b) *Vesika kasada.*

Informal Speech: The document is in the safe.

As Lewis points out, *-dir* functions as a copula in writing and formal speech (as in 12a), but it is ordinarily omitted in informal speech (as in 12b).⁴ The use of *-dir* as a copula in informal speech will usually have the effect of emphasizing the predicate, although such emphasis can actually render the predicate less certain, as can be seen from the translation of (12b) which uses the adverb *surely*. The Turkish enclitic *-dir* has a wide range of uses which are not comparable to the Lak assertive,⁵ but the essential dichotomy — formal:neutral/informal:emphatic — is analogous.

I would suggest, therefore, that the two different uses of the Lak assertive are reconcilable into a single meaning, viz., 'objective assertion.' Formal contexts are ordinarily supposed to be objective, and so the use of a form that specifies the predicate as an objective fact will be contextually neutral. Informal contexts, however, usually imply some degree of subjectivity, and so the use of a marked objective form will be contextually emphatic, i.e., emphasis on objectification (vs. personal opinion) is neutral in a formal context but emphatic in an informal one.

This in turn leads to my choice of the term *assertive* over, e.g., *confirmative* or *affirmative*. I eschewed the term *affirmative* (Russian *utverditel'nyj*) in my earlier work on Lak due to the fact that in English this term is often opposed to *negative* and *interrogative* as one of the three basic types of sentences, and I still hold to that view. In my work on Balkan languages (e.g., Friedman 1978), I have used the term *confirmative* to refer to screeves in which the speaker always personally

vouches for the truth of the predicate. Such forms are characterized, among other things, by the fact that they can only refer to events which have already occurred (i.e., are past) and can thus be truly vouched for as well as merely believed in. The Lak assertive occurs not only in the past and present but also, e.g., in the future (*čičinṭissara*) and conditional (*čičinṭissanija*, *čičajssanija*). Also, the common (plain, unmarked) forms to which the Lak assertives are opposed do not carry nuances of nonconfirmativity (e.g., reportedness) such as occur in the nonconfirmative past forms of the Balkans. Such nuances in Lak are carried by various analytic and syntactic constructions, e.g., those in *tar*, *unukkar*, *xxaj ur*, and the present gerund with *-ussa -ur*. The term *assertive*, which is one of the possible translations of the Georgian *mḡicebiti*, best captures the meaning of the Lak finite synthetic forms based on the participle in *-ss[a]-*. Just as in English, an assertion can be objective or aggressive (consider, e.g., the term *assertiveness training*), so also in Lak the assertive can be objective in formal writing or speech but aggressive, or emphatic, in informal speech.

This Lak phenomenon is extremely important to a broader understanding of status categories, i.e., categories involving a personal evaluation of the narrated event. Despite the obvious similarities between assertive and confirmative meanings, the two are clearly different gradations on a continuum of “positive personal involvement,” while nonconfirmative, unwitnessed, reported, etc., continue in the direction of “negative personal involvement.”

University of North Carolina, Chapel Hill

Notes

⁰I wish to thank the Library of the Georgian Academy of Arts and Sciences, which has kindly sent me materials essential to my study of Lak. I also wish to thank the Faculty Research Council of the University of North Carolina, Chapel Hill, which has financially supported my studies in Lak. In addition, I wish to thank Ms. Eleonora Magomedova and Ms. Füsün Leventoğlu for their helpful judgments on Lak and Turkish data, respectively, in an earlier version of this article.

¹Interestingly enough, Murqilinskij (1980:108-111) does not distinguish between assertive and nonassertive screeves in his section on tense in his Lak school grammar.

²Murkelinskij uses *opredelenno* only in future assertive screeves, which are beyond the defined scope of this article. Nonetheless, an example has been included in order to give a fuller picture of the nuances carried by assertive forms.

³This pair of forms is problematic in many respects, including the accuracy of the label *imperfect*, and I will treat them in a later work. In the remainder of this article, I

will restrict my attention to the forms labeled *present* and *past* (*čičajssar*, *čivčussar*, etc.).

⁴There are certain contexts in informal speech where *-dir* is not or cannot be omitted, but these are not relevant to this discussion (see Lewis 1967:97).

⁵See Lewis (1967:97-98, 139-41; also Friedman 1978:112-13). The analogies are potentially quite far-reaching, but they must be the subject of a separate study.

References

- Burčuladze, G. T. 1979. "Voprosy stanovlenija ličnogo sprjaženija v lakskom jazyke." *Iberiul-kavkasiuri enatmecnierebis keliḡdeuli* 6:176-246.
- Friedman, Victor A. 1978. "On the Semantic and Morphological Influence of Turkish on Balkan Slavic." *Papers from the Fourteenth Regional Meeting: Chicago Linguistic Society*; Donka Farkas, et al., eds. Chicago: Chicago Linguistic Society, 108-18.
- _____. 1984. "Status in the Lak Verbal System and Its Typological Significance." *Folia Slavica* 7:135-49.
- Lewis, Geoffrey L. 1967. *Turkish Grammar*. Oxford: Oxford University.
- Murkelinskij, G. B. 1971. *Grammatika lakskogo jazyka: I. Fonetika i morfologija*. Maxačkala: Dagestanskoe učebno-pedagogičeskoe izdatel'stvo.
- Murqılinskij, Xaži. 1980. *Lakku maz: 5-6 klassirtanssa učebnik*, 4th ed. Maxačkala: Dagučpedgiz.
- _____. 1981. *Lakku maz: 7-8 klassirtanssa učebnik*, 4th ed. Maxačkala: Dagučpedgiz.
- Xajdakov, S. M. 1953. *Značenie vremennyx form lakskogo jazyka*. (Dissertation summary). Moscow.
- _____. 1961. *Očerki po leksike lakskogo jazyka*. Moscow: Akademija nauk SSSR.
- _____. 1962. *Laksko-russkij slovar'*. Moscow: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovar'ej.
- Žirkov, L. I. 1955. *Lakskij jazyk*. Moscow: Akademija nauk SSSR.

Prehistoric Contacts Between Ossetic and Slavic

Zbigniew Gołąb

There is nowadays little doubt about close and prolonged prehistoric contacts between the North Iranian languages of the Scythians and Sarmatians on the one hand and Proto-Slavic on the other. Those contacts seem to have started in the 8th century B.C., i.e., when according to historical sources the Scythians invaded the Pontic steppes, and extended until at least the 5th century A.D., when the late Sarmatian tribe, the Alans, the direct ancestors of the Ossetians, still roamed the Pontic steppes.¹ Thus the whole period of contact lasted about twelve centuries.

Convincing linguistic evidence of those contacts are the Iranian loanwords in prehistoric Slavic that can be traced in the historical Slavic languages, and Iranian hydronyms in what is historically the Ukraine and Southern Russia.

Just for the sake of illustration the following examples can be quoted:

In the first place among Iranian loanwords in prehistoric Slavic one should mention an obvious semantic borrowing represented by the pair *bogъ* : *divъ* 'God': 'Demon' which obviously imitates the dualistic Iranian opposition *baga-* : *daēva-*, in this sense known only in Iranian and Slavic (cf., e.g., Lith. *diēvas* which preserves the primary IE meaning "God"); similarly another religious term *svętъ* 'sacred', imitating Iran. *spənta-* idem can be quoted. Its primary meaning has been preserved in Slavic names like *Svętopęlkъ*, *Svętoslavъ*, *Svętovitъ*, etc. where the first component means "strong, powerful." These and some other words represent a semantic adjustment of inherited IE elements in Slavic to their Iranian correspondences. But we also have many "material" loanwords from Iranian whose phonemic shape in Slavic proves their origin. These are first of all words with initial *x-* instead of IE *s-* according to the well-known Iranian phonemic change. E.g., *xotęti* (and its family) 'want' — Iran. *hant-* 'gelangen, gelangen lassen' (cf. Lith. with the preservation of the *s-*: *sintęti* 'sich entschliessen, denken, überlegen', (IE **sent-*; cf. Pokorny 908); *xvala* 'glory' — Iran. (Alanian) **xvālā-*, i.e., Iran. **xvāryā-*, cf. Av. *x^varənah-* 'Ruhm, Ruhmesglanz, etc', Sanskr. *svārṇara-* m. 'Lichtraum, Äther'; **xvorna* 'food' (this primary meaning is preserved in

South Slavic — e.g., S-C. *hrána*, and West Slavic, e.g., Kaszub. *charna* ‘fodder’) — Iran. *xvarəna*- ‘food, drink’ continued by Ossetic *xwar* ‘grain, barley’, etc.

The sum total of Iranian loanwords in prehistoric Slavic, both semantic and material, seems to reach thirty, which is not a negligible number. For comparison we can mention the number of Germanic loans in prehistoric Slavic, which amounts to forty-four words.

The hydronyms of Iranian origin in the Ukraine and Southern Russia have been the subject of numerous studies. Here it suffices to mention the names of three big rivers: *Don*, *Dniepr*, and *Dniestr*. The Iranian (or, to be precise, Ossetic) origin of the hydronym *Don* is obvious (cf. Oss. *don* ‘water, river’ from OIran. *dānu*- ‘river, big river’). The Iranian origin of *Dněpr̥* and *Dněstr̥* (these are the primary forms attested in Old Russian) is not so obvious and requires some explanatory remarks.

It has been commonly accepted that they represent primary compounds whose first component is OIran. *dānu*- ‘river’; the discussion (since Kretschmer) concerns the interpretation of the second components (cf. Vasmer I, 518). My contention is that for *Dněpr̥*, i.e., Proto-Slavic **Duncipr̥c(s)* we should start from OIran. **dān(u)-aipira*-² ‘river-upper’ = ‘upper river’: in such a situation this hydronym would primarily refer to the middle course of that river, i.e., its upper course from the standpoint of the Iranian tribes living along its lower course, called at the time of Herodotus Βορυσθενης (= OIran. *Vourusthāna*- ‘wide place’?). The transmission of the OIran. **Dān(u)-aipira*- to the prehistoric Slavs must have taken place through early Ossetic (= late Alanian, V-VII A.D.) because otherwise we would not understand the treatment of *Dān(u)*- as *Dñn*- in Slavic (cf. Oss. *don* where the *o* is closed!). And it must have happened before the monophthongization of diphthongs in Slavic, or at least during this process (400-600 A.D.), which would account for OIran. *ai* = *ě*. So the hydronym *Dněpr̥* is of Ossetic or late Alanian origin. It replaced the older Slavic name of this river, *Dunaj*, which has been preserved in Ukrainian and Byelorussian folklore.

Similarly the hydronym *Dněstr̥* should be derived from OIran. *Dān(u)-aēšra*-³ ‘river-swift’ through the same late Alanian/early Ossetic intermediary (for details see Goṭab, 1992).

So we are fully justified in asking whether there are any traces of prehistoric contacts with the Slavs in those contemporary Iranian languages which directly or indirectly continue the Scytho-Sarmatian group. As is well-known, Ossetic, spoken in the central part of the northern Caucasus, represents the only remnant of Alanian, a late Sarmatian dialect. Thus, etymological analysis of the Ossetic vocabulary may be revealing in this respect. To some extent the Iranian languages

and dialects located far to the east, Afghan (Pushtu) and the Pamirian group, which are considered descendants of the prehistoric Sakyan, the easternmost branch of primary Scythian, can also be taken into account. For it is possible that some western (i.e., Slavic) elements borrowed, for example, by the Pontic Scythians, spread in prehistoric times far to the east across the Eurasian steppes. Similar far-reaching migration of terms is frequently observed in the cultures and languages of large nomadic-pastoral groups. Of course, the question formulated above could only be answered by an expert in historical Iranian and comparative Indo-European linguistics. Fortunately, in recent times one outstanding Soviet Iranist of Ossetic origin, V. I. Abaev, has devoted to these problems a very revealing book, *Skifo-evropejskie izoglossy. Na styke vostoka i zapada*. As the title of the book indicates, the author is interested in the prehistoric isoglosses connecting Scythian (i.e., the old North Iranian branch of languages, including Scythian, Sarmatian, Alanian, and probably also primary Sakyan) and the IE languages of Europe. The author tries to show — and I think he succeeds — that Scythian had many old common isoglosses with Balto-Slavic, Germanic, and Italo-Celtic not shared by the other Iranian languages. These isoglosses, belonging mainly, though not exclusively, to the lexicon, seem to prove that the Scythian branch of the Iranian languages was autochthonous in Eastern Europe, i.e., it did not represent the “back wave” of the Iranian peoples. I would like to call attention to a paragraph in Abaev’s book, the conclusion of the chapter “North Iranian tribes — the original inhabitants of Eastern Europe” (Abaev 1965:121-24):

And just the opposite, everything falls into place once we admit that the Iranian element was present in southern Russia at least from the beginning of the second millennium B.C. Then the process that brought about the formation of Scythian-European isoglosses can be sketched in the following way. After the Iranian ethno-linguistic community in Southeastern Europe dissolved, one part of the tribes which had formed it moved south and east, to Media, Parthia, Persia, and Central Asia. Another part, the ancestors of the future Scythian tribes, remained in Europe and in the course of many centuries developed in contact with the peoples of Central and Eastern Europe area, including the future Italics and Tokharians. It was in this period that the identity of Scythian within the Iranian

language group was determined and numerous Scythian-European isoglosses originated.

For our purposes the analysis of old Slavic-Scythian lexical isoglosses established by Abaev acquires special significance. According to him, "The special Scythian-Slavic isoglosses far exceed in number and weight the special ties of Scythian with any other European language or language group" (Abaev 1965:135). Of course, Abaev speaks about isoglosses in most cases without deciding whether a given word, common, for example, to Scythian and Slavic, ultimately represents an element borrowed in prehistoric times from Slavic into Scythian, or simply continues a dialectal IE element of the contact area. To be sure, in most cases such a decision is even impossible. But I have scrutinized the list of those special Slavic-Scythian lexical isoglosses and I have been able to discover some which can be treated instead as prehistoric borrowings from Slavic into Scythian (i.e., modern Ossetic).

Incidentally, Iranian **tapara-* 'hatchet' (MPers. *tabrak*, NPers. *teber*, Kurd *tefer*, Beluchi *tapar*) should be considered a prehistoric Slavic loanword in Iranian, not vice versa (Slavic *toporъ* from *tepъ*, *te(p)ti* has an obvious derivational motivation within Slavic; cf. Vasmer 4:79), so we can expect that there will be more such old Slavic loanwords in Scythian and its descendants, first of all in Ossetic. This expectation seems to be fulfilled in the following examples, taken from the list of Slavic-Scythian isoglosses established by Abaev:

(1) Oss. *cæd(æ)* (**čata*) 'pair of oxen' ~ Proto-Slavic *četa*, e.g., Russ. dial. *četa bykov*, Church Slav. *četa*, *phálanx*, etc. (Abaev 1965:23).

(2) Oss. *cyt//citæ* (**či-θa-*) 'honor' ~ PSi. *čbstb* (**čbt-tb*) idem. The underlying verbal root **keit-* and its derivative in *-ti-*, e.g., Av. *čisti-* 'Denken, Einsicht', OIran. *čitti-* idem, are known in Aryan, but the Oss. word seems to have been affected semantically by Slavic *čbstb* (see Abaev 1965:31).

(3) Oss. *fæxt//fæstæ* 'stamping trough', from PSi. **pěstb*, cf. Russ. *pest* 'pestle', Pol. *piasta* 'hub, nave', and Lith. *piestà*, *piēstas* 'stamping trough'. Abaev here accepts a borrowing from Slavic (Abaev 1965:23).

(4) Oss. *qælæs//ġælæs* 'throat, voice' from PSi. **golsb* (?); Abaev (1965:19) suggests a borrowing from Church Slavic *glasb* (sic!) which makes no sense. Because of the *polnoglasie* perhaps it is instead from early East Slavic **golās̃b*.

(5) Afgh. *γumba* 'swelling' ~ PSi. *gqba*, e.g., OCS *gqba* 'sponge', Russ. *gubá*, Ukr. *húba* 'touchwood', etc. Abaev (1965:13) quotes only Lith. *gumbas* 'Auswuchs', but Slavic seems to be geographically closer and ultimately identical semantically: *gqba* 'any organic excrescence'.

(6) Oss. *xsyrf//æxsirf* (*xsirf* < **sirp-*) 'sickle', is undoubtedly borrowed from PSl. **sǫrpъ* (cf. Russ. *serp*, etc.), although Abaev (1965:13) simply treats it as a Scytho-Balto-Slavic isogloss. However, the treatment of **r* as *ir* (*ʁr*) is a characteristic Balto-Slavic innovation (IE **sṛpo-* > Balto-Slavic **sirpa-*), and furthermore, the term itself seems to be primarily unknown in Baltic (it is attested only in Latvian, where it may be a Slavic loanword). Ultimately the very fact of the preservation of the IE *s-* proves that the word cannot be an old IE dialectal element of the area continued in Ossetic.

(7) Oss. *xumætæg//xumætægi* 'simple, ordinary, insignificant' (about persons) ~ PSl. *kǫmetь* 'warrior, hero', Ukr. *kmit* 'free, prosperous peasant', etc. This example is highly controversial because Slav. *kǫmetь* seems to be a late borrowing from Vulgar Latin *comite* 'companion' (Vasmer 2:261), and in all Slavic languages it primarily denoted a dignified, not ordinary person (Abaev 1965:21).

(8) Oss. *kajyn* 'touch' ~ PSl. *kosnǫti*, *kasati* idem, e.g., Russ. *kósnut'sja*, *kasát'sja*. Lith. *kasýti*, *kàsti* 'scratch' is etymologically identical (IE **kes-*), but semantically remote (Abaev 1965:15).

(9) Oss. *k'yllaw//k'ullaw* 'hernia, rupture' ~ PSl. *kyla*, e.g., Church Slav. *kyla*, Russ. *kila* idem., Lith. *kūla* 'excrescence'. OHG *hōla*, Old Icelandic *haull* 'hernia, rupture' show a different vocalism (Abaev 1965:19).

(10) Oss. *lænk//læncæ* (**lanka-//*lankja-*) 'gully, valley' ~ PSl. **lǫka*, e.g., Russ. *luká* 'meadow in a river-bend', Pol. *łąka* 'meadow', etc.; and Lith. *lankà* 'Flusswiese; Tal', Tokh. *lenke* 'small valley'. There are no Aryan correspondences (Abaev 1965:17). Since the noun has obvious derivational motivation in Slavic, namely: **lęko*, **lękti* 'to bend' ≥ **lǫka*, **lǫkъ* 'bend', etc., and in Baltic as well (see Trautmann 1970:159-60), it is rather a borrowing from PSl.

(11) Oss. *mæcyn* 'wallow' (in a fluid), e.g., in the curse *Dæ tuǵy fæmæcaj!* 'May you wallow in your own blood!' ~ PSl. *močiti*, e.g., Russ. *močit'* 'wet', etc. If it is restricted to cursing, then it may be a borrowing (Abaev 1965:16).

(12) Oss. *būræ-mæz* 'miraculous glue', precisely 'yellow ointment', where the second component equals PSl. *mazъ* (from the verb *mazati* 'smear'), e.g., Russ. *maz'*, Pol. *maż*, etc. There are no Aryan correspondences. In my opinion it is a borrowing from PSl. (Abaev 1965:18).

(13) Oss. *mūt//met* 'snow' ~ Slav. nouns from the verb *metq*, *mesti* 'sweep', e.g., Russ. dial. *za-met* 'blizzard', Pol. *za-mieć* idem, Rumanian (from Slavic) *omete* 'snow', etc. It is a possible loanword (Abaev 1965:31).

(14) Afgh. *mēšta* 'dwelling place' ~ PSl. *město*, e.g., OCS *město* 'Ort, Platz, Stelle', Russ. *město* idem, etc. Because of Av. *maēθana-* 'dwelling place' (IE **meit-*) and the Balto-Slavic archetype **maista-* from the same verbal root (Trautmann 1970:185, s.v. *mintō-* 'wohne') I consider Afgh. *mēšta* an old Slavic loanword in the Scythian group (Abaev 1965:13).

(15) Sakian *pā'sa* (**parsa-*) 'pig' ~ Balto-Slavic **parša-*; cf. Lith. *pařsas*, PSl. **porse*, e.g., S-Cr. *prāse*, Pol. *prosię*, Russ. *porosēnok* idem, etc. The word may represent an old borrowing from Balto-Slavic or early Proto-Slavic (!) parallel to the borrowing of the same noun by the Finnic languages (e.g., Finnish *porsas*, etc.). As is well-known, the pig is a typical animal of sedentary farmers unknown to the pastoral nomads (Abaev 1965:12).

(16) Afgh. *pūnda* (**pāntā-*) 'heel' ~ PSl. *pęta*, e.g., OCS *pęta*, Russ. *pjată*, S-Cr. *pęta* idem. Lith. *pęntis* is formally more remote because of its *i*-stem. Perhaps, like *mēšta*, this is also an old borrowing from early Proto-Slavic (through Sakian), although there is a difference in the quantity of the root vowel (Abaev 1965:13).

(17) Oss v. *stæn* (regularly from **scæn-*) 'male dog', used only in the expression *stænmæ cæwyn*, literally 'go to a male dog', referring to a bitch in heat ~ PSl. **ščen-* (**sken-*), attested in Russ. *ščenók*, plur. *ščenjáta*; ORuss. *ščenja*, Gen. *-jate*, *ščenъcb*; Church Slav. *štenę*, Gen. *-ęte*; Pol. *szczenię*, Gen. *-ęcia*; S-Cr. *štène*, Gen. *-eta*, plur. *štenci*, etc., everywhere with the meaning 'puppy, young dog', undoubtedly conditioned by the diminutive character of the suffixes *-ęt-* and *-ъcb-*. The Oss. word has correspondences in Pamirian dialects: Bakh. *skan*, *skanok* 'puppy, young dog', which undoubtedly represents an earlier stage of PSl. *ščen-*, i.e., **sken-*. The word may be an old borrowing from early PSl. if we accept its derivation within Balto-Slavic from the verbal root **sken-*, cf. Latv. *skanēt* 'sound', Pol. *szczękać* (**sken-k-*) 'jangle, rattle' (Vasmer 4:502; Abaev 1965:14, 21-22).

(18) Oss. *tajyn* 'melt' ~ PSl. *tajati* idem, e.g., OCS *tájati* (*tajq*, *-ęši*), Russ. *tajat'*, etc. There is an exact correspondence, i.e., a stem in *-j-*, only in Slavic (Abaev 1965:15).

(19) Oss. (Dig.) *tærfæ* 'valley' ~ PSl. **torpǝ?*, e.g., Church Slav. *trapǝ* 'fovea, pitfall'. The Slavic word is thus semantically quite different; in addition, the noun is restricted to South Slavic (Skok 3:491). There is also Lith. *tárpas* 'Zwischenraum, Lücke, Kluft', and Tokh. *tarp* 'swamp, pond' (Abaev 1965:17).

(20) Oss. (Dig.) *tillæg* '[grain] crop, harvest' from PSl. *tblo* 'soil', e.g., OCS *tblo* 'Boden', etc. Abaev (1965:24) quotes only Germanic correspondences, e.g. Eng. *till* (<*tilljan*), but he does not realize that the Germanic verb is based upon a Proto-Slavic loanword in Germanic: PSl.

tblo > Germc. **tila*, from which the denominal verb **til-jan* > *tilljan* was derived. PSl. *tblo* represents a regular treatment of IE **telo-* with the \emptyset -vocalism of the root, i.e., **tlo-*, cf. Lat. *tellūs*, *-ūris* f. 'earth'; cf. also Lith. *tilės* f. pl., but with a different meaning 'boards at the bottom of a boat' (Pokorny 1959,1:1061). The form of the root in Ossetic indicates a borrowing from Proto-Slavic, which in the case of such a basic agricultural term would not be surprising.

(21) Oss. *ulæn* 'wool' (**ulnæ*) from PSl. **vblna* or early East Slavic **vblna*, e.g., OCS *vlbna*, S-Cr. *vūna*, Pol. *wetna*, etc., idem; Av. *varāmi-* 'wool' is formally far remote. It is rather surprising that the name of such an important product of the pastoral economy was borrowed from Slavic, but the borrowing may have been conditioned by barter trade (Abaev 1965:22).

(22) Oss. *wældæf* 'air', a compound of *wæl* 'top' (Russ. *verx*) and *tæf* 'spirit' (Russ. *dux*), an exact *calque linguistique* of PSl. (?) *vǝzduxǝ*, e.g., OCS *vǝzduxǝ* 'Luft, Hauch', Russ. *vǝzdux*, etc. (Abaev 1965:31).

(23) Oss. *wīs//wes* 'twig' (**vaiša-*) ~ PSl. **věxa*, e.g., Russ. *věxa*, Ukr. *vixá*, Pol. *wiecha*, etc. The basic meaning in Slavic seems to be 'stick with a straw bundle on top.' There are also some Germanic correspondences not mentioned by Abaev, e.g., Norw. *veis* f. (**vaisō-*) 'Stengel'. Since in other Iranian languages the parallel derivatives from the same verbal root **uei-* 'wind' show the suffix *-ti-*, this word may be an old borrowing from Slavic, before the early PSl. passage of **uoisā* > **uoixā* (Abaev 1965:20-21).

(24) Afgh. *wraža* (**bruša*) 'flea' ~ Balto-Slavic **blusā-*, i.e., PSl. **blǝxa*, e.g., Russ. *bloxá*, etc. This correspondence has already been quoted by Zaliznjak (1962:38) who has Afgh. *wraža* (sic!) (Abaev 1965:13).

(25) Afgh. *zanai* (regularly from **zǝna-ka-*) 'grain' (Russ. *zernó*) ~ PSl. **zbrno*, e.g., OCS *zbrno*, etc.; Lith. *žirnis* 'pea' (sic!). An old Scythian-European semantic correspondence, but perhaps ultimately influenced by early Proto-Slavic (Abaev 1965:13).

(26) Oss. *zīvæg* (**zaivaka-*) 'lazy, idler' from PSl. *zǝvati* 'yawn' with the Iran. suffix *-aka-*. The verb is well attested in Slavic, e.g., Russ. *zévát*, Pol. *ziewać*, etc., but unknown in Iranian (Abaev 1965:20).

I have supplemented the above list of Slavic-Scythian lexical isoglosses (according to Abaev) with some additional facts and remarks, which help us to distinguish among these isoglosses words representing probable Proto-Slavic loanwords in Scythian (i.e., historically mainly in Ossetic). As such I consider the following:

1. Oss. *cæd(æ)* from PSl. *četa*;
2. Oss. *fæstæ* from PSl. **pěstǫ// *pěsta*;
3. Oss. *ǵælæs* from PSl. **golǫ/*golǫs (?)*;
4. Oss. *xsyrf* from PSl. **sǫrpǫ*;
5. Oss. *lænk* from PSl. **lǫka*;
6. Oss. *mæcyn* from PSl. **močiti*;
7. Oss. *-mæz* from PSl. **mazǫ*;
8. Oss. *met* from PSl. **-metǫ*;
9. Afgh. *měšta* from PSl. **město*;
10. Sak. *pā'sa* (**parsa-*) from early PSl. **porsǫ*;
11. Oss. *stæn* and Pamir. (Vakh.) *skæn* from early PSl. **skenǫ/*ščenǫ*;
12. Oss. *tillæg* from PSl. **tǫlo*;
13. Oss. *ulæn* (**ulnæ*) from PSl. **vǫlna/*vǫlna*;
14. Oss. *wes* from early PSl. **uǫisǫ*, later **věxa*;
15. Afgh. *zanai* (**zǫnaka-*), semantically influenced by PSl. **zǫrno (?)*;
16. Oss. *zīvæg* from PSl. **zěvati*.

As we see, there are sixteen words⁴ which with different degrees of probability can be considered as old borrowings from Proto-Slavic into Scythian. Among them agricultural terms (or terms connected directly and indirectly with agriculture) are the most characteristic: *cæd(æ)*, *fæstæ*, *xsyrf*, *pā'sa*, *tillæg*, *zanai*. Of course, not all of these borrowings stem from the same epoch. Those which are attested in Afghan and Pamirian dialects, as probably inherited from Sakian, must be the oldest, i.e., they were borrowed before the middle of the first millennium B.C. (Abaev 1965:12-13). But the Slavic loanwords in Ossetic also represent a varied chronology, e.g., *wes* from early PSl. **uǫisǫ* must be older than *ulæn* from East Slavic **vǫlna*, etc. So the very fact of the chronological differences between individual loanwords seems to support the idea of a long period of close ethnolinguistic contacts between the prehistoric Slavic and Scytho-Sarmatian tribes. That period undoubtedly started much earlier than the eighth century B.C., the traditional date of the alleged Scythian invasion of the Pontic region. Abaev argues convincingly against this traditional view and accepts the autochthony of the Scythians in Eastern Europe. If we agree with him, then the contacts between the linguistically separate Proto-Slavs and Iranian Scythians (already crystallized as separate ethnolinguistic entities) must have started at least about 1000 B.C. So the whole period of prehistoric ethnolinguistic contacts and exchange between the Slavs and the Scytho-Sarmatian tribes lasted 1500 years (until the sixth century A.D.?). It would be incomprehensible if the linguistic traces of such a long period of contact were insignificant on both sides. The lexical facts I have presented and discussed here seem to corroborate the hypothesis that this contact was

really fertile (productive) in linguistic-cultural exchange. Here I would like to repeat, with some additions, Abaev's remarks (1965:142-43) on from whom the Scythians learned agriculture. The Scythians unquestionably learned agriculture from the natives of Europe. As Abaev states, in Ossetic there is only one agricultural term, *jæw* 'millet', which was inherited from Common Iranian and Aryan. Other agricultural terms have many old and obvious correspondences in the European languages. Abaev quotes the following: *xsyrf* 'sickle', with correspondence in Slavic: *fsir* 'ear of grain', with correspondence in Germanic; *fsonz* 'yoke', with correspondences in Baltic, Germanic, and probably Latin; *stīvz* 'plug connecting a yoke with its shaft', with correspondences in Germanic; *fæxt/fæstæ* 'mortar', with correspondences in Slavic and Baltic; *tillæg* 'harvest, crop', with correspondences in Germanic (this statement, as I have shown, is wrong; *tillæg* is a borrowing from PSi. *tblō*). Among these six agricultural terms three are of Slavic origin. We should also add here *cæd(æ)* from PSi. *četa*. Thus, the Proto-Slavs seem to have been the people from whom the nomadic Scythian shepherds learned agriculture.

University of Chicago

Notes

¹According to T. Sulimirski (1970:164-65) in the 4th-5th centuries A.D. the Alans were concentrated in Bessarabia along the Prut river. This region is known in the historical sources of the time as *Alania* and the river as *Alanus fluvius*. In this connection, the Alanic origin of the Bessarabian city name *Iasi* (ORuss. *Jasskyi tǫrg*) from the ethnicon (*Iasi*) is obvious.

Another important piece of evidence for prolonged contact between prehistoric (Eastern) Slavs and the Sarmatian tribes is the ethnicon *Antes*, *Antae*, etc. used commonly by the 6th century Latin and Greek sources for the Eastern Slavs between the Dniepr and Dniestr rivers. As has been convincingly argued by Vasmer (1964: s.v. *Vjātiči* (sic!)) this ethnicon was most probably borrowed by the Greeks and Romans from the Pontic Iranians (Sarmatian Alans) and represents an Iranian name given by the latter to their northwestern neighbors dwelling along the northwestern frontier, as the very etymology of the name indicates. For we are dealing here with a derivative of Aryan *ānta*- 'Ende, Grenze, Rand', i.e., either with the adjective represented by OIran. *antya*- 'am Rande befindlich, am Ende befindlich' or **ānta*- 'belonging to a border, etc.', so the meaning of *Antes*, *Antae* would be 'frontiersmen' (cf. the historical parallel in the primary semantics of *Ukraincy*). Since the Eastern Slavs themselves have not preserved any vestiges of this ethnicon, I think that speculations (e.g., those by T. Sulimirski) about primary Sarmatian ethnicity of the *Antes* etc., who were allegedly only gradually slavized, is pointless.

But the contacts between the early historical Eastern Slavs and the Alans (particularly the *Asi, Jasi* tribe) were not broken in the 5th century A.D., i.e., when the Huns dominated the steppe regions of Eastern Europe (from the Sea of Azov to the Lower Danube). As is known, the eastern branch of the Alans survived between the Don river and the Caucasus, and it is there that the Old Russian chronicles mention several expeditions against them by the Kievan princes, up until the beginning of the 13th century (Gaglojti 1966:189-202).

²**aipira-*, an adjective representing IE **epi-ro-*, cf. Av. *aipi* (preposition) 'über-hin, bei', primarily 'auf-darauf', e.g., Albanian *epërë* 'oben befindlich' (Pokorny 1959, 1:323-24).

³**aēšra-*, an adjective representing IE **eis-ro-*, cf. Av. *aēš-* (IE **eis-*) 'sich in eilige Bewegung setzen' (Pokorny 1959, 1:299).

To the list taken from Abaev (1965) one could add words to be probably found in Abaev 1958-. But the dictionary is not completed as yet and we cannot check the list of Slavic forms included in it. Nevertheless, even a very perfunctory look at the entry words may provide us with some interesting facts. For example, the word *suxæ* (Digor) 'dry' (*suxæ xwæsxæ* 'dry hay', *suxi mæjæ* 'month of drought [=August?]) represents a regular Slavic development of IE **sausō-* (cf. Lith. *sausas*, Gr. *αὔρος*) opposing Iran. *haus-*, attested in Oss. *xūs*, *xōysk*'. To be sure, Abaev does not accept here a prehistoric Slavic loanword, and attempts to interpret the case as an example of a so-called "perekrestnaja isoglossa" ("intersecting isogloss"), but the whole concept of such isoglosses presented by him in a separate article (1966:247-63) seems to be too vague to be theoretically convincing.

References

- Abaev, V. I. 1958-79. *Istoriko-ëtimologičeskij slovar' osetinskogo jazyka*, 1-3. Leningrad.
- _____. 1965. *Skifo-evropejskie izoglossy: Na styke vostoka i zapada*. Moscow.
- _____. 1966. "O perekrestnyx izoglossax." *Ëtimologija*, pp. 247-63.
- Gaglojti, Ju. S. 1966. *Alany i voprosy ètnogeneza osetin*. Tbilisi.
- Gołąb, Zbigniew. 1992. *The Origins of the Slavs: A Linguist's View*. Columbus, Ohio.
- Pokorny, J. 1959-69. *Indogermanisches etymologisches Wörterbuch*, 1-2. Bern, Munich.
- Skok, P. 1971-74. *Etimologijski rječnik hrvatskoga ili srpskoga jezika*, 1-4. Zagreb.
- Sulimirski, T. 1970. *The Sarmatians*. New York, Washington.
- Trautmann, R. 1970. *Baltisch-slavisches Wörterbuch*, 2nd ed. Göttingen.
- Vasmer = Fasmer, M. 1964-73. *Ëtimologičeskij slovar' russkogo jazyka*, 1-4. Moscow.
- Zaliznjak, A. A. 1962. "Problemy slavjano-iranskix jazykovyx otnošenij drevnejšego perioda." *Voprosy jazykoznanija* No. 6:28-45.

On The History of Relative Clauses in Georgian¹

Alice C. Harris

1. **Introduction.** Old Georgian had a single type of relative clause: a postnominal clause initiated with a relative pronoun. Modern Georgian, on the other hand, has an unusual variety of relative clause types. The purpose of this paper is to examine how a single strategy developed into the variety found in Georgian today.

In the account below, I follow the typology of Maxwell (1979), with terminological changes suggested in Maxwell (1982). In general, Maxwell's typology differs from those of Comrie (1981), Givón (1979), Keenan (1985), Keenan and Comrie (1977) only slightly. In the present paper, **head (noun)** refers to the noun in the main clause which is modified by the relative clause. **Relativized nominal** refers to a nominal in the relative clause which is coreferential to the head.

2. Synchronic Analysis.

2.1. *Old Georgian.* Old Georgian possessed only a single type of relative clause.² The relative clause in Old Georgian follows the noun it modifies, the head noun. The relative clause is introduced by a relative pronoun or adverb. While a number of relative pronouns and adverbs were used, including *vin* 'who' and *sada* 'where', it is only the most frequent, *romel* 'which', that we shall be interested in here. Generally, the case of the relative pronoun is determined by its grammatical role in the relative clause, as illustrated in (1). Its number agrees with that of the head noun.

- (1) *da aha k'ac-i ert-i, romel-sa saxeli xerkua zekey*
and behold man-NOM one-NOM *which-DAT* name call *Zaccheus*
'And behold a man, *whose name was Zaccheus.*'
(Xanm. L 19:2, cited by Ziziguri 1973:222.)

Aside from variations due to differences of case and number, the pronoun *romel* occurred in several forms. In the nominative singular we find the following, among others:

- | | | |
|------------------|--------------------|---------------------|
| (2) <i>romel</i> | <i>romel-igi</i> | <i>romel-igi-ca</i> |
| <i>romel-i</i> | <i>romel-i-igi</i> | <i>romel-i-ca</i> |

In Old Georgian nouns, the regular nominative case in *-i* alternates with the archaic absolute (stem) form. The two forms have the same privileges of occurrence, except that the stem form is more likely to be found in indefinite environments.³ The pair *romel/romel-i*, however, does not adhere to this general pattern of distribution, and *romel* is found in environments where it is clearly definite, as illustrated by (3), (4) and (5).

- (3) *da jač'v-i igi, romel edva kedsa missa*
and chain-NOM DEF *which it/lie/to.her neck-DAT her*
'the chain *which lay on her neck*'
(Šuš. IX, 24; cited by Ziziguri 1973:217)
- (4) *enmanuel, romel ars gamotargmanebit: čuen tana γmerti*
Emanuel.NOM *which it/be translation.INST us with god*
'Emanuel, *which is in translation: God with us*'
Mt 1: 23 Ad; cited by Ziziguri 1973:218)
- (5) *avazak'-ni igi, romel jur-cumul iq'vnes mis tana*
thief-PL.NOM DEF *which cross-beaten they/be him with*
'the thieves *which were crucified with him*'
(Mt 27: 44 Ad; cited by Ziziguri 1973:218)

Thus, *romel* and *romeli* are interchangeable in Old Georgian.

The forms *romel-igi* and *romel-i-igi* are best viewed diachronically as relative pronouns plus a demonstrative. The form *igi* functioned as a demonstrative meaning 'that', as a definite article 'the', and as a third person singular personal pronoun '(that) he/she/it'. The forms *ese* and *ege* 'this' had a similar range of functions and are likewise found with *romel(i)*. However when *igi*, *ege*, and *ese* occur with *romel(i)*, they differ from the demonstratives, article, and personal pronoun in that they do not vary for case and number, as illustrated in (6).

- (6) *k'ac-sa mas, romel-sa-igi qeli ganqmel hedga*
man-DAT the *which-DAT-PRT hand withered it/be/to.him*
'to the man *which had a withered hand*'
(Haem. Mark 3: 3; cited by Ziziguri 1973:230)

In (6), the relative pronoun itself, *romel-sa*, is in the dative. The form *igi*, enclitic to *romel-sa* in (6), is the same as the nominative case of the demonstrative, not the dative. Thus, in this particular function *igi* is indeclinable and may be considered a relative particle synchronically in Old Georgian.

We may observe that among languages of the world it is fairly common to have a demonstrative used together with a relative word. For example, in early Old English we find such a usage.

- (7) *þæt is se Abraham se him engla god naman niwan asceop.*
 that is that Abraham that to.him angels' God name new made
 'That is the Abraham for whom the God of the angels made a new name.'

(Exodus 380-1; cited by Visser 1963:522; Romaine 1982:57)

Here *se* functions as a demonstrative agreeing with Abraham, while *him* bears the dative case as required by the syntax of the relative clause.

At all stages of Georgian, the relative pronoun may occur with the invariant particle *-ca*.

- (8) *γirso mama-o, romel-man-ca locv-ita šenita*
 worthy father-VOC which-NAR-PRT prayer-INST your

momanič'os zleva-y ešmak'tay
he/give/it/to.me power-NOM devils

'worthy father, *who by your prayers will give to me the power of devils*'
 (Xanzt. ey, v. 42; cited by Ziziguri 1973:233)

In Old Georgian the particle *-c(a)* was optional, while in the modern language it is obligatory.

In summary, there is no significant difference in distribution among the several forms in (2).

2.2. *Middle and Modern Georgian.* In Modern Georgian we find a great diversity of relative clause types. The relative clause construction found in Old Georgian remains in the modern language, as illustrated in (9).

- (9) *čit-i, romel-i-c bambi-dan gak'etebuli-a*
 calico-NOM which-NOM-PRT cotton-from made-it/be
 'calico, which is made from cotton'

Morphologically this type is more restricted than in Old Georgian, since (a) the particle *-c(a)* is now obligatory, (b) the particle *-igi* is no longer permitted, and (c) there is no longer an alternation between *romel* and *romeli*. Syntactically (9) is like Old Georgian (1)-(6). In what follows, this is termed the **(postnominal) relative pronoun strategy**.

All of the remaining types, introduced to the language in historical times, make use of a particle *rom/ro* 'that'. The first of the new types may be referred to as the **postnominal gap strategy**; it is illustrated in (10).

- (10) *xalxi k'areb-tan axlos ro idga, aq'aq'anda.*
 people(SG) doors-at close that he/sit he/clap
 'The people *who sat close by the doors* began to clap.'

- (11) *ert-i matgani tma-ši rom bandi akvs čac'nuli*
 one-NOM them *hair-in that band he/has/it tied*
 'one of them *who has a band tied in his hair*'
 (Examples (10) and (11) cited by Vogt 1971:51.)

In a gapped relative clause, the relative nominal is not overtly represented in its clause. In the postnominal gap type, the relative clause follows its head. The head noun bears the case required by its function in the main clause. The subordinate clause always contains the particle *rom/ro*, which occurs somewhere before the verb of its clause, usually in second position.

In addition to postnominal relative clauses, Modern Georgian has relative clause types that precede the head noun. We may consider first the **prenominal gap strategy**, illustrated in (12).

- (12) *šen-gan ro mivi-γeb, im pul-it me*
you-from that I/receive/it that money-INST I/NOM
gadavixdi val-s.
I/pay/it debt-DAT
 'I will pay off the debt with that money *which I receive from you.*'
 (Tschenkéli 1958:203)⁴

Like clauses produced with the postnominal gap strategy, this clause type is characterized by a relative clause lacking an overt representation of relative nominal and containing the particle *ro(m)*. The prenominal gap type generally precedes not only the head noun, but the entire main clause, as in this example.

The final major relative clause type is the **(prenominal) non-reduction strategy**, in which the relative nominal is represented as a noun, rather than being "reduced" to a pronoun or gap. According to the structure of the main clause, three subtypes can be distinguished.

- (13) *mindā, Betania-ši rom k'olmeurnoba-a, is vnaxo.*
I/want/it Betania-in that collective-it/is it/NOM I/see/it
 'I want to see the collective-farm *that is in Betania.*' (Cited by Vogt 1971:51.)

In (13) the main clause contains a pronoun coreferential with the relative nominal, *k'olmeurnoba*. In the twelfth century example (14), on the other hand, there is no resumptive pronoun.

- (14) *me rom tkven-gan movisminen c'q'aloba-ni, mediada.*
I/NAR that you-from I/hear/them mercy-PL/NOM it/overwhelms/me
 'The grace *which I have heard from you* overwhelms me.'⁵
 (Vepxist'q'aosani 134, 3; cited by Ertelišvili 1962:153.)

In a third subtype the whole nominal is repeated in the main clause, as in (15).

- (15) *Durmišxan-s Alget-ze rom c'iskvili eč'ira*, is
Durmishxan-DAT Alget-on that mill he/have/it that
c'iskvili...
mill
'the mill which Durmishxan had on Alget'
 (Sizv. 74, 1; cited by Ertelišvili 1962:153.)

Although (14) and (15) appear in Middle Georgian literary works, these types do not occur in the standard modern language.

Thus, Middle and Modern Georgian together have an unusually large number of strategies for the formation of relative clauses. Examples include most of the types included in most typologies, including all of the subtypes of the (prenominal) non-reduction strategy.

3. Diachronic Analysis. While Old Georgian relativized only with relative pronouns, Middle Georgian used three strategies. The use of relative pronouns remained productive, and the postnominal gap strategy began to be used frequently. The non-reduction strategy, while it was used at this period, appears to be less common. In Modern Georgian we find the three attested from Middle Georgian, as well as the prenominal gap strategy.

Studies by Ertelišvili (1963) and Ziziguri (1973) have established, among other things, (i) that relatives other than the postnominal relative pronoun type developed in historical times, (ii) that the strategies developed in the three general stages outlined in the preceding paragraph, though exact orders have not been established, (iii) that the general purpose subordinating particle *rom/ro* develops via *rome* from *romel*, the relative pronoun 'which' in Old Georgian. It remains to investigate how the full variety of syntactic structures found in the modern language developed.

3.1. Development of the Postnominal Gap Strategy. In Old Georgian, the relative pronoun filled three functions at once: (a) It signalled subordination, specifically the onset of a relative clause. (b) It represented the relative nominal. (c) Its case indicated the surface grammatical relation of the relative nominal in its clause.

In Old Georgian, the standard was for the relative pronoun to bear the case appropriate to its function in the relative clause. However, some examples show so-called "progressive case attraction", a construction in which the relative pronoun shows the case of its head noun, rather than

the case determined by the syntax of its own clause (Aronson 1972; Dondua 1967).

- (16) *xiq'os me vin tkuen-gan-i k'ac-i, romel-i*
 he/be INDEF someone you-from-NOM man-NOM *which-NOM*
txxovdes ze-y misi p'ur-sa
he/ask/him son-NOM his bread-DAT
 '(if) there be some man among you, *whom his son asked for bread*'
 (Xanm. Mt. 7: 9; cited by Dondua 1967:23.)⁶

Without case attraction, the expected form of the relative clause in (16) would be *romel-sa* (DAT) *txxovdes zey misi p'ursa*. In such examples, *romeli* fills functions (a) and (b) above, but not (c).

In the examples of "progressive attraction" known to me in Old Georgian, the relative pronoun is always "attracted to" a head in the nominative and itself has the form *romel* or *romeli*. I suggest that this form, not motivated by the syntax of the relative clause, as was the norm, was interpreted as an invariant particle, just as *igi*, *ege* and *ese* were when enclitic to the relative pronoun (see above). I suggest that *romel(i)* was interpreted as a relative clause subordinator, much like English *that*. Once *romel(i)* was understood this way, it ceased to fill function (b), and the relative noun was no longer represented. The clause would then have to be understood as having a gap. The relative clause in a sentence like (16) originally had a surface structure in which the relative pronoun was interpreted as the indirect object, as schematized in (17); it later was interpreted as having a structure like (18), where 'Ø' represents the gap, and *romel(i)* is a subordinating particle.

- (17) *romel-sa txxovdes ze-y misi p'ur-sa*
 which-DAT he/ask/it son-NOM his bread-DAT

- (18) *romel(i) [txxovdes ze-y misi Ø p'ur-sa]*

After (17) was reanalyzed as (18), *romeli* no longer filled functions (b) or (c), but only served to mark subordination, the onset of the relative clause.

While this hypothesis provides a plausible account of the development of the postnominal gap strategy, instances of progressive case attraction are quite rare in Old Georgian, and it is likely that another, more important, mechanism also played a role in the development of this new relativization strategy.

Though the absolute form (stem form) of the nominative case was still found in indefinite environments in Old Georgian, by Middle Georgian stem forms of nouns had become quite rare. The stem form of the

relative pronoun, *romel* 'which', however, did occur. Since *romel-i*, with the overt case marker, or *romel-ni*, with the nominative plural suffix, was now the norm, the old stem form could be understood as an indeclinable particle. In (19) we see an example.

- (19) *igi č'abuk'-ni, romel dges*
 those youth-PL/NOM *which-Ø they/stand*
 'those youths *who stand*'
 (Amiran-Darejaniani 72, 30; cited by Ertelišvili 1963:151.)

In Old Georgian it was not unusual to represent a nominative plural by the stem form,⁷ and it would have been natural for this usage to be continued into Middle Georgian. However, stem plurals fell out of use and the stem form was understood as not bearing a case. In this way *romel* was open to interpretation as an invariant subordinating particle, and the clause containing it would be understood as having a gap, rather than a relative pronoun. While (19) would originally have been related to the alternative phrasing, (20), it was later understood as having the structure in (21), where *romel* is a subordinating conjunction.

- (20) *romel-ni dges*
 which-PL/NOM they/stand

- (21) *romel [Ø dges]*

Originally in sentences like (19), *romel* had functions (a), (b), and (c); later it was understood as having function (a) alone, introducing a subordinate clause.

In the history of Georgian, we can witness the reduction of the particle *romel* to *rome*, then *rom*, and in many dialects to *ro* today. (22) gives an example of *rome* from Middle Georgian; (11) shows *rom* in the standard modern language; while (10) illustrates the popular *ro* form.

- (22) *igi samni k'ac-ni, rome lomta zeda sxdes*
 those three man-PL/NOM *that lion on they/sit*
 'those three men *that sit on a lion*'
 (Amiran-Darejaniani 24, 6; cited by Ziziguri 1973:256.)

In such examples it is clear that *rome* functions as a subordinating particle, rather than as a relative pronoun, since the form is invariant (cannot be declined).

The reduction of the particle *romel* seems to be due in part to a need to distinguish it from the pronoun *romel-i*, and in part to the existence of forces of erosion that apply – however sporadically – to liquids, vowels, and nasal consonants in the Kartvelian languages (Žyent'i 1953). The

exemption of the pronoun *romel-i* from these same processes is explained in part by the fact that it is “protected” by *-i* and the other endings it has (case suffixes or the particle *-c*), and in part by the need to differentiate the pronoun and the particle.

The problem of word order remains: According to the proposal made here, the clause-initial relative pronoun *romel* was reinterpreted as a subordinating particle or conjunction, reduced by phonological attrition to *rom*. An example which seems to be a direct continuant of this word order is (23).

- (23) *es ori ojaxi iq'o, rom mamxnevebda.*
 this two family it/be that it/make.brave/me
 ‘There were these two families *that made me brave*.’
 (Ak’ak’i 460; cited by Ziziguri 1973:256.)

Ziziguri (1973:257) states that it is entirely natural to have *rom* in clause-initial or clause-second position in the modern language. Nevertheless, examples with *rom* later in the clause, such as (10) and (11), appear to occur more frequently. Furthermore, consultants with whom I worked on this **do not permit** *rom* to be clause initial in at least some examples in the postnominal gap construction; it must be in “second” position. For example, my consultants judged (24a) grammatical, with *rom* in second position; but (24b), with *rom* in clause-initial position was deemed ungrammatical.

- (24a) *es k’aci, c’igni rom p’ropesor-s mout’ana*
 this man book that professor-DAT he/bring/it/to.him
 ‘this man who brought the book to the professor’
 (b)**es k’aci, rom p’ropesors c’igni mout’ana*⁸

The position that *rom* occupies in Standard Modern Georgian is not strictly **second**; according to my consultants, *rom* may occur anywhere between the first constituent of the clause and the verb. For example, (24c), with the same meaning as (24a), was judged grammatical, though *rom* is actually the third constituent in the clause.

- (24c) *es k’aci, p’ropesors c’igni rom mout’ana*

Given the grammaticality of (23), it seems reasonable to assume that the ban against clause-initial *rom* is a new development, perhaps not (yet) shared by all speakers.⁹

3.2. Development of the Non-Reduction Strategy. In Old Georgian we find also “regressive case attraction”, a construction in which the head noun agrees in case with the relative pronoun, instead of bearing the case

determined by the function of the head in the main clause. Examples are (25) and (26).

- (25) sit'q'ua-sa, *romel-sa get'q'ude me tkuen*, sulī ars.
word-DAT *which-DAT I/tell/it/to.you I you soul it/be*
'the word *which I say to you* is spirit'
(John 6: 63 Ad; cited by Dondua 1967:24.)

- (26) q'ovel-i, *romel-i xitxovs*, miiγis.
all-NOM *which-NOM he/ask/it he/receive/it*
'All *which request* will receive.'
(Haem. Mt. 7: 8; cited by Dondua 1967:25.)

The expected form of (26), without attraction, would be as in (27).

- (27) q'ovel-man, *romel-i xitxovs*, miiγis.

I suggest that structures like (26), more common than "progressive case attraction", were reanalyzed by rebracketing. That is, the head noun, marked with the case appropriate to the relative clause counter the grammatical norms, was reinterpreted as belonging to that clause rather than to the main clause. According to my proposal, the structure of (26) would originally have been as in (28), and later as in (29).

- (28) q'ovel-i, [*romel-i xitxovs*], miiγis.
all [*which ask*] receive

- (29) [q'ovel-i *romel-i xitxovs*], miiγis.

With attrition, we would get the variant in (30).

- (30) [q'ovel-i *rom xitxovs*], miiγis.

The development proposed here explains the origin of the clause-second position for *rom* in the non-reduction construction. The relative pronoun *romel* had obligatorily occupied clause-initial position. Absorbing into the subordinate clause the head, which had preceded the relative clause, would necessarily push *rom(el)* to second position, as schematized in the transition from (28) to (29). We may assume that later the constraints on "second" position were weakened to allow the subordinator to occur anywhere between the first word of the relative clause and the verb (see section 3.1).

In this way, sentences like (26) can be seen as leading to examples like (14). However, this structure, without the resumptive pronoun, is apparently possible in Modern Georgian only with a meaning such as 'whoever'. The immediate ancestor to the non-reduction construction

found in the modern language, (13), was instead the construction in example (31).

- (31) *mama-y igi šeni, romel-i xedavs daparul-sa,*
 father-NOM the your *which*-NOM *he/see secret*-DAT
mogagos man šen cxadad.
 he/reward/you he/NAR you openly
 ‘your father, *which sees in secret*, will reward you openly.’
 (Mt 6: 6 Ad; cited by Dondua 1967:26.)

In this example, the head, *mama-y igi šeni* ‘your father’, is in the case required by the lower clause; the pronoun *man* ‘he’ is coreferential with the head and bears the case required by the syntax of the main clause. It is this structure which seems to have given rise to Modern Georgian examples like (13), where a resumptive pronoun in the main clause follows the relative clause.

Thus, examples with the absolute form and “case attraction”, both of which were at this time dying out, came to be seen as unmotivated for some speakers, leading to their reinterpretation as having the invariant form *romel*. In both the postnominal gap structure and the non-reduction structure, the form *romel* was eroded, becoming *rome*, *rom*, and often *ro*. Declined forms of the relative pronoun, such as *romel-s*, could coexist with them because the declined forms were synchronically motivated from the point of view of the syntactic requirements of the lower clause.

3.3. *The Naturalness of “Case Attraction”.* Both hypotheses developed above assume that “case attraction” is a natural process that needs no special explanation. That is, both depend upon the idea that it is natural for the head noun to agree in some instances with the case of the relative pronoun or vice versa, due to their proximity and coreference. The idea that such agreement is natural and requires no special explanation is supported by the fact that it occurs in many languages. For example, it is found optionally in Latin, Ancient Greek, and Old English. It has been reported that regressive case assimilation is obligatory in Nandi, a Nilo-Saharan language (Keenan and Comrie 1977).

Dondua (1967), on the other hand, has taken the view that case attraction is something unnatural that requires special explanation. He has suggested that regressive case attraction develops because in examples like (25) and (26) the head noun was **originally** part of the relative clause. He suggests that after the relative pronoun came to be viewed as the beginning of the clause, the head noun was in the “wrong” case from the point of view of the syntax of the main clause, of which it was now a member. Thus, in many ways, Dondua’s hypothesis is the opposite of mine.

It may be noted that Dondua's proposal will not account for progressive case assimilation, wherein the relative pronoun agrees with the case of the head that precedes it. Nor does he give any account for how the non-reduction strategy might have developed, if not in this way. These might be seen as reasons for rejecting Dondua's proposals.

However, I think Dondua's view is not necessarily inconsistent with the proposal developed here. It is not impossible that the head noun did start out in the relative clause, with *romeli* following it as a demonstrative, in the way suggested by Dondua. Then the head may have come to be viewed as part of the main clause, and in most examples it bore the case necessary for membership in the main clause. However, the construction with regressive case attraction would have continued to exist as a possibility and would have been grammaticalized. Once the clause-initial position of *romel* had become obligatory, as in Old Georgian, it was possible to reverse the process, with the head again being viewed as part of the relative clause, this time with *romel* no longer seen as a demonstrative or even a relative pronoun, but as an invariant form. Thus, I think the two views are not necessarily inconsistent, though they at first appear to be.

3.4. *Development of the Prenominal Gap Strategy.* The development of the prenominal gap strategy is perhaps the least clear. I assume that it developed by analogy to postnominal gap, on the one hand, and prenominal non-reduction, on the other.

The development of the prenominal gap strategy, like that of the (prenominal) non-reduction type, was motivated in part by a need to put relative clauses, like all other adnominal modifiers, before the head. In Old Georgian, all modifiers, including adjectives, quantifiers, determiners, and relative clauses followed the head noun either in unmarked order (adjectives, determiners, relative clauses) or optionally (quantifiers). In Modern Georgian the unmarked order is for all modifiers, with the exception of the postnominal relative clauses described above, to precede the head (Harris 1985b). Thus, the development of prenominal relatives is part of a more general change in the order within nominal phrases.

Universally, the gap strategy may occur pre- or postnominally. Typological studies cited earlier suggest that this is not true of the relative pronoun and non-reduced types. We may assume that the development of the prenominal gap strategy in Georgian is an **extension**. As defined informally in Harris (1985a: Ch. 15), an extension is the removal of conditions on a rule. In this instance, the condition that was removed was the language-particular constraint specifying that the gap strategy could occur only postnominally.

4. Conclusion. The history of Georgian is unusually rich in changes of strategy for the formation of relative clauses. We have here an example of a postnominal relative pronoun strategy spawning both the postnominal gap type and several subtypes of the prenominal non-reduction strategy. To these, in turn, was added the prenominal gap strategy. A study of these developments in Georgian not only pinpoints the mechanisms of these changes within this language, but, within the context of general linguistics, adds to our understanding of how such changes may come about.

Vanderbilt University

Notes

¹This material is based upon work supported by the National Science Foundation under Grant No. BNS-8419143; it draws also on research on Modern Georgian in 1974-75, supported by the International Research and Exchange Board under the Graduate Student/Young Faculty Exchange with the Soviet Union.

²The Old Georgian relative clause was of a single type, whether we follow the classificatory schema of Givón (1979), that of Keenan and Comrie (1977), or that of Maxwell (1979).

³This reflects its origin as the indefinite member of the opposition \emptyset /*i*. In spite of the fact that their functions are nearly identical, these two forms are traditionally treated as two cases. I have argued that they are forms of a single case (Harris 1985a: Ch. 4).

⁴Tschenkéli put the comma after *pulit*; I have placed it as required by the intonation used by my consultants.

⁵This translation is not the same as that of any of the published translations into English, but is based on Ertelišvili's paraphrase in Modern Georgian.

⁶Dondua seems to suggest that this is the only such example, but additional ones are given in Ertelišvili 1963:151.

⁷An example is *romelta esxnen col* (1 Cor 7: 29) 'those who had wives' where 'wives' is expressed with the stem form, *col*, but triggers plural agreement; see Harris 1985a:212.

⁸The particle *rom* is used in many other kinds of clauses, including ones expressing 'if' and 'when'. An interpretation of that sort might be possible for this sentence, were it not for the semantics of the main clause.

⁹Another obvious possibility is that the grammaticality of (23) is related to the fact that it is extraposed.

References

- Aronson, Howard I. 1972. "Some notes on relative clauses in Georgian." *The Chicago which hunt: Papers from the relative clause festival*, 136-43. CLS parasession volume.
- Comrie, Bernard. 1981. *Language universals and linguistic typology*. Chicago: University of Chicago Press.
- Dondua, K'arpez. 1967. "Mimartebiti nacvalsaxelisa da misamarti sit'q'vis urtiertobisatvis zvel kartulši." [The relationship between the demonstrative pronoun and the correlative word in Old Georgian.] *Rčeuli našromebi*, I, 20-29. Tbilisi: Mecniereba.
- Ziziguri, Šota. 1973. *K'avširebi kartul enaši*. [Conjunctions in the Georgian language.] Tbilisi: Universit'et'i.
- Ertelišvili, P. 1963. *Rtuli c'inadadebis ist'oriisatvis kartulši*, 1: *Hip'ot'aksis sak'itxebi*. [On the history of the complex sentence in Georgian, 1: Questions of hypotaxis.] Tbilisi: Universit'et'i.
- Givón, Talmy. 1979. "From discourse to syntax: Grammar as a processing strategy." *Discourse and syntax (Syntax and semantics 12)*, ed. by Talmy Givón, 81-112.
- Harris, Alice C. 1985a. *Diachronic syntax: The Kartvelian case*. New York: Academic Press.
- _____. 1985b. "Word order reconstruction in Kartvelian." Paper presented to the Annual Meeting of the Linguistic Society of America.
- Keenan, Edward L. 1985. "Relative clauses." *Language typology and syntactic description*, II, ed. by Timothy Shopen, 141-170. Cambridge: Cambridge University Press.
- _____ and Bernard Comrie. 1977. "Noun phrase accessibility and universal grammar." *Linguistic Inquiry* 8:63-99.
- Maxwell, Daniel. 1979. "Strategies of relativization and NP accessibility." *Language* 55:352-371.
- _____. 1982. "Implications of NP accessibility for diachronic syntax." *Folia Linguistica Historica* 3:135-152.
- Romaine, Suzanne. 1982. *Socio-historical linguistics*. Cambridge: Cambridge University Press.
- Tschenkéli, Kita. 1958. *Einführung in die georgische Sprache*, I. Zürich: Amirani.
- Visser, F.T. 1963. *An historical syntax of the English language*, I. Leiden: E.J. Brill.
- Vogt, Hans. 1971. *Grammaire de la langue géorgienne*. Oslo: Universitetsforlaget.
- Žyent'i, Sergi. 1953. "Kartvelur enata dialekt'ebis šesc'avlis zogierti sak'itxi I. St'alinis saenatmecniero šromebis šukze." [Some questions of the study of the dialects of Kartvelian languages I. In light of Stalin's linguistic work.] *Tbilisis universit'et'is šromebi* 49:129-42.

Notes on Auxiliary Verbs in Tsova-Tush (Batsbi)⁰

Dee Ann Holisky

In contrast to the extended discussions of auxiliaries one finds in literature on English, they are scarcely mentioned in descriptions of the languages of the Caucasus. Descriptions of Tsova-Tush are no exception: Schiefner (1859:71) noted in passing one verb which could be considered an auxiliary, and the Tsova-Tush dictionary (Kadagije and Kadagije 1984) has an entry for that verb and one other auxiliary.¹ One might thus conclude, erroneously, that such a category was rare or nonexistent in this language.

The purpose of this paper is to present evidence for a subclass of auxiliary verbs in Tsova-Tush, and to discuss some of their morphological and syntactic properties. In addition, it is hoped that this discussion will stimulate the investigation of such phenomena in other languages of the Caucasus.

In the first section of this paper, I will briefly describe how class agreement and case government work in Tsova-Tush clauses which have only one verb, as a necessary background to understanding the patterns found with auxiliaries. In the second section one type of construction with two verbs (one finite, one non-finite) will be discussed, in order to differentiate the syntactic patterns exhibited by this construction from those found with auxiliaries. In the third section, the patterning of two-verb constructions with auxiliary verbs will be defined and illustrated. Then, morphological and syntactic properties of five different auxiliary verbs will be discussed.

1. Patterns of concord in Tsova-Tush. In Tsova-Tush there are three kinds of concord between a verb and its arguments: case, person, and class. (Person and class agreement also show, secondarily, number.) Concord of case and class will be the main focus here; person markers are occasionally important because they provide some evidence about case.

Case. Verbs may occur with one, two, or three arguments, which appear in one of five cases: Nominative, Ergative, Dative, Contact, or Allative.² The number of arguments and the cases they must appear in is determined by the individual verb; there are at least eleven different valence patterns in Tsova-Tush.

For example, the single argument of a one-place verb may be in the Nominative, the Ergative, or the Dative; the two arguments of a two-place verb may be in the Ergative and the Nominative, the Dative and the Nominative, the Ergative and the Dative, and so on. There are three one-place patterns, five two-place patterns, and three three-place patterns.

Two different three-person patterns are illustrated in (1).³

- (1a) k'natev mezoblen burt qaʔ-b-ieⁿ
 boy-ERG neighbor-DAT ball-NOM hit-throwing-AOR-CL-ball
 'The boy hit the neighbor with a ball.'
- (b) k'natev ph"arax t'ot' lah"iⁿ
 boy-ERG dog-CON hand-NOM touch-AOR
 'The boy touched the dog.'

Thus, in constructions with more than one verb, the case of an argument can often provide clear evidence about which verb it is governed by. This kind of evidence will be drawn on below.

This kind of evidence is not always available, however, because there is frequently no overt expression of the arguments of the verb. Arguments which are understood from context and are not contrastive or emphatic are not expressed (cf. the Georgian rule of Unemphatic Pronoun Drop in Harris 1981: 32-33).

In the absence of expressed arguments, the case of the person markers can provide evidence about the construction. The relevant facts about agreement with person are these: A verb may show explicit agreement with a first or second person subject or object; there are no overt markers indicating third person. Person marking is achieved by means of verbal suffixes, nearly identical in form to the corresponding personal pronouns.

There are two sets of person suffixes: Nominative forms are used to mark agreement with an argument in the Nominative case and Ergative forms to mark agreement with an argument in the Ergative. (2a) shows person agreement with a subject in the Nominative, (2b) with a subject in the Ergative, and (2c) with a direct object in the Nominative.

- (2a) v-ek'-v-ieⁿsO 'I trembled.'
 tremble-AOR-1SG-NOM
 CL-SG-M
- (b) kh"eklas 'I'm getting ready.'
 prepare-PRES-1SG-ERG

- (c) v-ik'eⁿsO 'They took me in.'
 take-AOR-1SG-NOM
 CL-SG-M (A17.1)

Class. There are eight noun classes in Tsova-Tush; some (but not all) verbs show class agreement with one of their arguments by means of prefixed class markers. (In many lexically-derived verbs, the class marker will be in the middle of the verb, since it is the verbal suffix which takes a "prefixed" class marker, e.g., in the verb in (1a).) Class markers are set off from the verb and, where relevant, the argument with which it agrees is indicated in the verb gloss.

A verb usually shows class agreement with (the referent of) an argument in the Nominative case.⁴ Thus, in (1a), the verb shows agreement with *burt* 'ball', which requires a marker *-b* in the singular. Other examples are given below; in (3a) *v* shows agreement with *vasO* and in (3b) *j* shows agreement with *jasO*.

- (3a) nanas vasO v-ik'eⁿ
 mother-ERG brother-NOM took-away-AOR-CL-M
 'Mother took brother away.'

- (b) nanas jasO j-ik'eⁿ
 mother-ERG sister took-away-AOR-CL-F
 'Mother took sister away.'

Evidence using class agreement will be drawn on below, as a means of distinguishing different syntactic patterns.

2. **Finite verb as main verb.** Tsova-Tush has a host of nonfinite verbal forms, including participles, verbal nouns, converbs, and infinitives. Here we will consider only constructions with infinitives, since it is necessary to distinguish two different kinds of constructions in which the infinitive is used.

In one construction, the subject matter of this section, we have a two-clause construction, where both the infinitive and the finite verb govern separate arguments. The finite verb can be shown to be the head of the construction. In the other infinitival construction, discussed in the following section, the infinitive and finite verb always govern the same arguments, and it is the infinitive which is the head.

To turn, then, to constructions of the first type, the use of infinitives is licensed by a number of different verbs, including 'want', 'try', 'help', 'begin', 'end'. Infinitives are also used in purpose clauses with many other verbs. Some examples are given in (4) and (5). (The infinitive

marker is *-an* but via regular phonological processes most word-final nasals nasalize the preceding vowel and delete.)

- (4a) h²amina[?] le[?]E ait':veⁿ xil'aⁿ
 everyone-DAT want-PRES happy become-INF
 'Everyone wants to become happy.' (Kadagije and Kadagije 1984:31)

- (b) duq cad-b-aliⁿ baq:axui o kor-v-ah'aⁿ
 much try-AOR-CL- leader-PL- 3SG-NOM capture-INF-CL-
 PL-M NOM SG-M
 'The leaders tried hard to capture him.' (G21.5)

- (c) joh² bader d-apx-d-aⁿ j-ol-j-aliⁿ
 girl-NOM child-NOM undress-INF-CL-child begin-AOR-CL-SG-F
 'The girl began to undress the child.'

- (5a) monadir h²uilnO v-axeⁿ b[?]orc' b-[?]evaⁿ
 hunter-NOM woods- go-AOR- wolf-NOM kill-INF-CL-wolf
 2ndDIRN CL-SG-M
 'The hunter went to the woods to kill a wolf.'

- (b) dani[?] maiqI j-aq'aⁿ xabžeⁿ
 all-NOM bread-NOM eat-INF-CL-bread sit-down-AOR
 'Everyone sat down to eat.'

- (c) c'igI ecaⁿ v-[?]eviⁿ seⁿ dad
 blood-NOM take-INF kill-AOR-CL-SG-M my father-NOM
 '(They) killed my father to take blood.' (A18.6)

As the examples indicate, these are all same-subject constructions: The finite verb and the infinitive are always understood to have the same subject. In (4b), for instance, 'leaders' is understood as the subject of both 'try' and 'kill'; in (5b) 'everyone' is subject of both 'sit' and 'eat'.

Moreover, the subject is expressed only once; it is always the **finite** verb which determines its case. In (4a), 'everyone' is governed by 'want' and not 'become', which requires a Nominative subject. In (4c), 'girl' is governed by the finite 'begin' and not the infinitive 'undress', which takes an Ergative subject. In (5a) 'hunter' is governed by 'go' and not by 'kill', which requires an Ergative subject. For this reason, the finite verb in this construction is referred to as the 'main verb.'

Although there are two Nominative arguments in many of the examples, neither verb could alone could support two Nominatives. That is, the Nominatives must be analyzed as arguments of different verbs. In (5b) the first Nominative ('all') is subject of 'sat', while the

second ('bread') is direct object of 'eat'. In (5c), each verb governs a different direct object ('blood' is object of 'take' and 'father' is object of 'kill').

As for class agreement, the two verbs in this construction frequently show agreement with different arguments. In (4c), for instance, 'begin' agrees with 'girl', while 'undress' agrees with 'child'.

In summary, except for the restriction to same subject, the finite verb and infinitive can have independent arguments, each verb retaining its own valence and class-agreement. This construction clearly consists of two clauses, a main clause and a reduced infinitival complement.

3. Finite verb as auxiliary verb. There are other verbs which appear in construction with an infinitive, however, which show a distinctly different pattern, as shown in (6) and (7).

- (6a) je, inc k'nativ lat'daⁿ later ...
and now boy-PL-ERG add-INF AUX-IMPF
'And now the boys kept on adding ...' (A1 4.6)
- (b) b^əark'i latE xil'aⁿ j-aq:aⁿ
eye-PL-NOM AUX-PRES become-INF big
'The eyes get big regularly.' (A1 26.11)
- (7a) h"on qenix nan j-ec'E ^əe-j-agaⁿ
you-DATnext-to mother-NOM AUX-PRES-CL-SG-F sit-INF-CL-SG-F
'Mother should be sitting next to you.'
- (b) oqus xenex co j-ec'rasO j-exk'aⁿ
3-SG-ERG tree-CON not AUX-IMPF- tie-INF-CL-
1SG-NOM-CL-SG-F SG-F
'He shouldn't have tied me to the tree.'
- (c) gornk'en t'q'uih" dah" v-ec'er v-ik'aⁿ
hill-DAT behind PVB AUX-IMPF-CL-SG-M take-INF-CL-SG-M
'They were to take him behind a hill' (A163.2)

In these examples, which are same-subject constructions like those discussed above, it is the infinitive which determines the case of the subject. For example, in (6a) the subject is in the Ergative, as required by 'add'; in (6b) it is Nominative, as required by 'become'. Though the examples in (6) have the same finite verb, their subjects are in different cases.

In fact, the infinitive determines the case not only of the subject, but of all arguments in the unit. That is, unlike the construction discussed in 2., there are never arguments which couldn't be supported by the

infinitive alone (were it to appear in finite form). It is not possible for these verbs to govern different direct objects (as was the case in (5c)).

The infinitive also determines the case of the person markers, though it is the finite verb which is inflected for person. In (8a) the finite verb shows agreement with a first person Nominative subject (as required by 'become') and in (8b) it shows agreement with an Ergative subject (as required by 'go').⁵ (An example of a finite verb showing agreement with the Direct Object was given in (7b).)

(8a) vorl'-barl' šarluⁿ xil'aⁿ j-olrasO
 seven-eight year-GEN be-INF AUX-IMPF-1SG-NOM-CL-SG-F
 'I was probably seven or eight years old.' (A149.16)

(b) moh" v-ec'es v-axan-ainO?
 how AUX-PRES-1SG-ERG-CL-SG-M go-INF-said-CL-SG-M
 'How should I go, he said?' (A127.16)

As for class agreement, the infinitive and finite verb (if they show class-agreement) always agree with the same argument. In (7a) both verbs agree with 'mother'; in (7c) both show agreement with a singular male (the unexpressed direct object). In this construction it is not possible for the verbs to agree with different arguments.

In short, this construction does not consist of separate verbs with independent arguments, representing two clauses. It is a monoclausal verbal unit, which governs one valence pattern and one class-agreement pattern.

The concord patterns of the infinitive determine the pattern of the unit and the lexical meaning of the infinitive determines the "selectional restrictions" for the unit (which NPs, for example would be semantically appropriate as subject or object). For these reasons, the infinitive is considered the head, or main verb.

Since the finite verb is inflected for screeve⁶ and person and contributes aspectual or modal modifications to the lexical meaning of the main verb, it is called an "auxiliary."⁷

4. Auxiliary verbs in Tsova-Tush. Given this general characterization of an auxiliary-main verb unit, I want to turn now to the individual auxiliary verbs in Tsova-Tush. I am aware of five, three with aspectual meaning and two with modal meaning. Three of these verbs are very common, occurring frequently in texts; *d-all* and *d-olO* are rather infrequent. The auxiliaries are listed in (9). (They are cited in the Present, except for *latiⁿ*, which is in the Aorist.)

(9) Aspectual Auxiliaries

- (a) *latE* regularly, repeatedly
- (b) *latiⁿ* begin to do regularly, repeatedly
- (c) *d-all* be about to, on the verge of

Modal Auxiliaries

- (d) *d-ec'E* should, must
- (e) *d-olO* probably, possibly

All five auxiliaries have defective paradigms. None has nonfinite forms, viz., masdar, infinitive, or participle. None occurs in the full range of tense-aspect forms found in Tsova-Tush, listed in (10).⁸

(10) Present Series of Screeves (Imperfective)

Present

Imperfect

Imperfect Reported

Aorist Series (Perfective)

Aorist

Pluperfect

First Aorist Reported

Second Aorist Reported

Except for *latiⁿ*, the auxiliaries are restricted to the screeves of the Present Series. The auxiliary *d-olO* is even more restricted, lacking an Imperfect Reported form as well.⁹ *Latiⁿ* occurs only in two screeves of the Aorist (the Aorist and the First Aorist Reported).

In the sections to follow, I will discuss each auxiliary separately, considering its meaning and restrictions on the kinds of verbs it occurs with. In addition, I will explore the degree to which the auxiliary-infinitive unit forms a cohesive constituent by considering factors like the relative ordering of auxiliary and infinitive, and the degree to which other elements (preverbs, particles, etc.) can occur between them.

It is interesting to observe that the different auxiliaries exhibit varying degrees of cohesion with the infinitive, ranging from instances where the auxiliary is a relatively independent word which appears in many syntactic positions to those where it is perceived as a suffix of the verb.

4.1 *latE* 'regularly, repeatedly'. A very frequently encountered auxiliary is *latE*, one of the two auxiliaries cited in Kadagije and Kadagije (1984). It contributes a durative aspectual meaning to the clause, highlighting the regular, or usual, nature of the state or action depicted by the verb. Sometimes it conveys repetition. Depending on context, it is

The auxiliary *latE* occurs with verbs denoting states as well as actions, with transitive and intransitive verbs. It does not occur, however, with verb roots which are grammatically perfective,¹⁰ as shown in (14). In

(14a) and (14b), the imperfective form of the root must be used, and the verb in (14c) which has no corresponding imperfective, is unacceptable with *latE*.

- (14a) PERF *lat'aⁿ latE
 fight-PERF-INF AUX-PRES
 'He fights regularly.'
 IMPERF let'aⁿ latE
- (b) PERF *o st'ak' sart'aⁿ latE
 DEM man-NOM curse-IMPF-INF AUX-PRES
 'This man curses all the time.'
 IMPERF o st'ak' sert'aⁿ latE
- (c) PERF *o k'nat šarn egaⁿ latE
 only DEM boy-NOM disappear-INF AUX-PRES
 'This boy disappears regularly.'

This restriction is motivated by the semantic incompatibility of the two categories: the perfective encodes a single, completed action (Holisky 1985:454), while *latE*, emphasizing as it does the repetitive or habitual nature of the action, encodes an action which is necessarily not a single one.

As for word order, the most frequent order in texts was infinitive followed by auxiliary, though the reverse was occasionally found (in 3 of 12 examples). A native-speaker consultant strongly preferred infinitive-auxiliary order.

In those less common examples where the auxiliary precedes the infinitive it is possible to have preverbs and negatives occur between the auxiliary and following infinitive, as in (15). The two do not form a particularly close unit.

- (15) obi lamnax co latE ču tivaⁿ
 they-NOM mountain-PL-CON not AUX-PRES PVB rest-INF
 'They don't regularly rest (vacation) in the mountains.'

However, when the preferred order occurs, the infinitive and auxiliary clearly form a syntactically cohesive unit. It is not possible to have preverbs, negatives, or particles (conjunctive, question, or intensive particles) between the infinitive and auxiliary, as shown in (16).

- (16a)*co lepsaⁿ h"alO later (with preverb)
 not dry-INF PVB AUX-IMPF
 'It regularly wouldn't dry.'

- (b) *j-elaⁿ co latE (with negative)
 laugh-INF not AUX-PRES
 'She doesn't laugh regularly.'
- (c) *lexane latE (with conjunctive clitic -e)
 look-for-INF-& AUX-PRES
 'He is always looking for it and...'

Further evidence for cohesion comes from the fact that the auxiliary *latE* is perceived as a suffix: My consultant, who helped transcribe Tsova-Tush narratives, writes it immediately following the infinitive, and commented on my writing it as a separate word, "I think these should be written together." In texts recorded by Rusudan Gagaa it is also written as a suffix.

The etymology of this auxiliary is unclear. Although there are two homophonous verbs ('to fight' and 'to be stuck') a more likely source would be *lat:ar* 'to be standing'. Nichols (Forthcoming) reports a similar, though restricted use of the Ingush cognate.¹¹

4.2 *latiⁿ* 'begin to do regularly, repeatedly'. A second very common auxiliary, no doubt related to *latE*, is *latiⁿ*, which contributes inceptive meaning to the clause. Specifically, it pinpoints the beginning of a durative action or activity (more rarely, a state). Examples are given in (17).

- (17a) *lexaⁿ* *latiⁿ* *c'er-k'itxv qetinO xalx*
 look-for-INF AUX-AOR literate people-NOM (A1 32.14)
 'They began looking for people who knew how to read and write.'
- (b) *zakairI* *latiⁿ* *sogO* *vuğaⁿ*
 Zacharia AUX-AOR 1SG-ALL call-INF (A1 49)
 'Zacharia began calling to me.'
- (c) *qeⁿ* *čuv-aixnes* *h"epsaⁿ* *latiⁿ* *sogO*
 then in go-AOR-1SG-ERG-& look-INF-IMPERFAUX-AOR 1SG-ALL
 'Then I went in and they began looking at me.' (A 19.1)
- (d) *qerl'aⁿ* *latiⁿsO*
 be-afraid-INF AUX-AOR-1SG-NOM
 'I began to be afraid (regularly).'

The main verb must depict something which will continue for a time, usually something recurring regularly. For example, (17d) means the subject began to be afraid on numerous occasions; it is not used to express the beginning of a single instance of being afraid.

The fact that the situation depicted by the main verb must recur may account for this restriction: *lati*ⁿ does not easily cooccur with many stative verbs, as shown in (18). Though it is easy to imagine being afraid on repeated occasions, once one knows someone or something, a state is normally viewed as enduring.

- (18a) *gak'vetil qetaⁿ latiⁿ
 lesson-NOM know-INF AUX-AOR
 'She began to know the lesson.'
- (b) *bader d-ep'c'aⁿ latiⁿ
 child-NOM know-INF-CL-child AUX-AOR
 'He began to know the child.'

Stative verbs (besides 'be afraid') which the consultant found acceptable with *lati*ⁿ include 'be cold' and 'be hungry', states which frequently recur.

*Lati*ⁿ, like *latE*, never occurs with perfective verbs; imperfective roots, if they exist, must be used. This restriction, a purely grammatical one, is explained by the semantic conflict between the perfective, which encodes a single action, and *lati*ⁿ, which requires a repetitive one. Examples are given in (19).

- (19a) PERF *v-at'aⁿ latiⁿ
 run-INF AUX-AOR
 'He began to run (away).'
- IMPERF it'aⁿ latiⁿ
 run-INF AUX-AOR
 'He began to run (around).'
- (b) PERF *dek'aⁿ latiⁿ
 fall down AUX-AOR
 'They began to fall down.'
- IMPERF ak'aⁿ latiⁿ
 fall down AUX-AOR
- (c) PERF *k'nat dah'' at'aⁿ latiⁿ
 only boy-NOM PVB fell-quiet-INF AUX-AOR
 'The boy began to fall silent.'

As for word order, as above, the most frequent order in texts was infinitive followed by auxiliary. The reverse order occurred only twice in fifteen examples.¹² Moreover, the consultant strongly preferred

infinitive-auxiliary order, most often rejecting the reverse, unless a negative preceeded the auxiliary.

The infinitive and auxiliary form a syntactically cohesive unit. When in unmarked order (infinitive followed by auxiliary), it is not possible to have words or particles occurring between them.

4.3 *d-all* 'be about to, be on the verge of'. A third Tsova-Tush auxiliary with aspectual meaning, *d-all*, is very rare. I've encountered it less than a half-dozen times in texts. This auxiliary emphasizes that an action or change of state is just about to take place; it can be glossed 'be about to', 'be on the verge of', as in (20).

- (20a) ph"arav j-ailrasO qah"ane, šarn j-ait'nas
 dog-ERG AUX-IMPF bite-INF- & self ran-off-AOR-
 1SG-NOM 1SG-ERG
 'The dog was about to bite me, so I ran off.'
- (b) kist'i mak b-ailI txon qetaⁿ
 kisti-NOM PVB AUX-PRES-CL-PL-M 1-PL-DAT attack-INF
 'The Kisti are about to attack us.'
- (c) učiⁿ j-alir xil'aⁿ
 black AUX-IMPF become-INF
 'It was on the verge of becoming dark.'

This auxiliary occurs only with perfective verb forms; it does not occur with nonperfectives of any kind, including statives, as illustrated in (21) and (22).

- (21) IMPERF *duixt'rev joh" j-alir deq:-j-aⁿ
 doctor-ERG girl-NOM Aux-IMPF examine-INF-
 CL-SG-F CL-SG-F
 'The doctor was about to be examining the girl.'
- PERF duixt'rev joh" j-alir daq:-j-aⁿ
 'The doctor was about to examine the girl.'
- (22) STATIVE *joh" j-ailI maic-j-aⁿ
 girl-NOM AUX-PRES-CL-SG-F be-hungry-INF-CL-SG-F
 'The girl is about to be hungry.'

The relative order in this case is inflexible and the reverse of that required by the two auxiliaries discussed above. *D-ailI* must precede the infinitive. Moreover, the auxiliary and following infinitive form a cohesive phrasal unit; they cannot be separated by other words or particles, except for a personal pronoun (20b) or the intensive particle (as in 23b).

- (23a) *duixt'rev j-alih"oE daq:-j-aⁿ (with conjunctive
 doctor-ERG AUX-PRES- examine-INF particle)
 2-SG-NOM-&

'The doctor is about to examine you and ...'

- (b) o h"alO v-alicI mac-v-alaⁿo
 3SG-NOM PVB AUX-PRES-I-CL-SG-M be-hungry-INF-CL-SG-M
 'He is really just about to become hungry.'

As further evidence for the constituency of the auxiliary-infinitive unit, preverbs (which usually immediately precede the verb which "requires" them), precede the auxiliary-infinitive unit, as in (24).

- (24) o dah" v-alir v-aš-v-alaⁿ
 3SG-NOM PVB AUX-IMPf-CL-SG-M get-shaven-INF-CL-SG-M
 'He was about to get shaven.'

Etymologically, *d-ailI* is probably related to the lexical formant (*dalar*), which forms intransitive inceptives when suffixed to verb stems. It is no doubt cognate to Ingush 'arrive' (Nichols Forthcoming).

4.4. *d-ec'E* 'should, must'. The most common modal auxiliary is *dec'E* 'should, must', noted in both Schiefner (1859:71) and Kadagije and Kadagije (1984). This very frequently occurring auxiliary seems to express a wide range of modal meanings from obligation to necessity. I know of no restrictions on the kind of verbs it occurs with. Examples are given in (25).

- (25a) lat'aⁿ b-ec'e-t-vai-ainO (A131.2)
 help-INF AUX-PRES-1PL-said
 'We should help him, they said.'
- (b) moh"e tamdes al'uicI, uist' d-ec'E xil'aⁿ (A143.2)
 how tamada-ERG say-PRES-I so AUX-PRES be-INF
 'Just as the tamada says, that's the way it should be.'
- (c) gornk'en t'q'uih" dah" v-ec'er v-ik'aⁿ (A163.2)
 hill-DAT behind PVB AUX-IMPf-CL-SG-M take-INF-CL-SG-M
 'They were to take him behind a hill.'

This auxiliary plus infinitive construction is relatively more flexible than the others. Though the clearly preferred order is auxiliary plus infinitive (occurring in 29 of 32 text examples), it is by no means fixed. The consultant, while preferring auxiliary followed by infinitive order, freely accepted the reverse.

There is less cohesion between the auxiliary and main verb as measured by possible occurrence of preverbs, negatives, and various particles between the two verbs, as in (26).

- (26a) dad v-ec'er ču tivaⁿ (with preverb)
 father-NOM AUX-IMPF-CL-SG-M PVB rest-INF
 'Father should have rested.'
- (b) co d-ec'ra č'ʔairkoⁿ al'aⁿ (with adverb)
 not AUX-IMPF-2SG-ERG suddenly say-INF
 'You shouldn't have said it suddenly.'

This auxiliary is clearly related to the common Dative-subject verb *d-ec'ar* 'want, love'. The widespread use of this verb as both main verb meaning 'want' and auxiliary verb meaning 'should', patterns identically to the Georgian verb *unda*. (Tsova-Tush speakers have been bilingual in Georgian for a long time.)

4.5 *d-olO* 'probably'. The final auxiliary to discuss is *d-olO*, which expresses possibility and probability. This modal is not common, but occurs with a wide range of verb types; there seem to be no restrictions with respect to co-occurrence.

- (27a) vorl'-barl' šarluⁿ xil'aⁿ j-olrasO (Al 49.16)
 seven-eight year-GEN be-INF AUX-IMPF-1SG-NOM
 'I was probably seven or eight years old.'
- (b) dah" daxaⁿ d-olor
 PVB drunk-INF AUX-IMPF
 'They probably got drunk.'
- (c) šukia ču toh"aⁿ j-olO
 Shukia PVB sleep-INF AUX-PRES-CL-SG-F
 'Shukia is probably sleeping.'

There are no possible variations in word order with this auxiliary. It always follows the infinitive. It forms a very cohesive syntactic unit with the infinitive; nothing may occur between them.

Like all other auxiliaries, this one occurs only in the Present Series of screeves, and in addition, it lacks the modal screeve of this series, the Imperfect Reported. This may be due to the fact that this auxiliary itself is already formally a modal (probably an older modal form of 'to be'), formed with the same suffix as is used in the Imperfect Reported, i.e., *-lo*.

4.6 Summary. In this paper it has been demonstrated that a category of auxiliary verbs has become established in Tsova-Tush, and that auxiliaries are distinguished syntactically from other verbs which occur

with infinitives. Formally, all of the auxiliaries have defective paradigms (occurring only in screeves of one series), and all lack nonfinite forms.

Semantically, the five auxiliaries discussed here have lost what lexical meaning they had and have taken on modal or aspectual meanings: obligation and probability for the modals, emphasis on duration, the begining of duration, or beginning of a perfective action or event for the aspectuals.

The auxiliaries have, at the same time, lost their own case and class agreement patterns. Although the auxiliary inflects for person, number, screeve, it is the main verb which determines the case pattern and selectional restrictions. In terms of valence and concord, the auxiliary and main verb function as a unit, with only one set of arguments.

They form a syntactic constituent with the infinitive which is unvarying to some extent. Moreover, although they are for phonological reasons separate words, the three auxiliaries which follow the infinitive are more closely connected to the infinitive than the others and show evidence of becoming suffixes. The main facts about order and possible insertion of material between the two verbs are given in (28).

(28)		Preferred order	Insertion of material	
(a)	latE	Inf + Aux	No	} resemble suffixes
(b)	lati ⁿ	Inf + Aux	No	
(c)	d-olO	Inf + Aux	No	
(d)	d-aill	Aux + Inf	Rarely	
(e)	d-ec'E	Aux + Inf	Yes	

Since there is not a similarly well-established category of auxiliary verbs in the related Chechen or Ingush, we can conclude that this is a Tsova-Tush innovation. In that light, the information presented here not only provides further data for the understanding of the category of auxiliary verbs from a universal point of view, it also gives evidence relevant to diachronic considerations.

George Mason University

Notes

⁰Tsova-Tush, the designation its native speakers prefer, is commonly known to linguists as Bats or Batsbi. It is spoken by about 3,000 people in the village of Zemo Alvani in eastern Georgia. The research reported here was supported by the International Research and Exchanges Board (summer 1983) and the National

Academy of Sciences (winter 1982-83). I express my appreciation to these and to the Linguistic Institute of the Georgian Academy of Sciences.

I am especially indebted to the many Tsova-Tush speakers who provided me with oral texts without which I would never have discovered this category. Special thanks are due to the family which hosted my stay in Alvani and to Elene Kadagije, who transcribed the texts and served as a Tsova-Tush-speaking consultant and colleague.

¹Schiefner (1859:71) cites *d-ec'E* 'should', while Kadagije and Kadagije (1984) cite this and *latE* 'regularly, repeatedly' as well. (See Note 3 on transcription; these verbs are cited in the Present Tense.)

²It is possible, but rare, for a Genitive to be governed by the verb:

- (i) vux leic'I h"e"n?
 what hurt-PRES you-GEN
 'Where do you hurt?'

This occurs with a limited number of verbs, usually in constructions which are idiomatic or archaic (e.g., curses).

³Examples are from texts collected in Zemo Alvani in January and May of 1983 (coded with Al), from texts collected by Rusudan Gagua (G), or from elicitation. In the transcription, nasalized vowels are indicated with superscript *n*, glottalized consonants with raised comma, final "reduced" vowels with a capital letter (their exact phonetic quality is unclear); /l'/ represents a voiceless lateral fricative, /j/ a palatal glide, /x/ and /ǧ/ velar fricatives, /q/ a voiceless uvular stop, /h"/ and /ʕ/ pharyngeal fricatives. Sequences of two vowels represent falling diphthongs.

Abbreviations in glosses: PRES for Present Tense, IMPF for Imperfect, AOR for Aorist, PERFective, IMPERFective, PVB for preverb, NOMinative, ERGgative, DATive, GENitive, CONtact (Georgian *gamsc'valviti*), ALLative, DIRN for Directional, CLass, DEMonstrative, Female, Male, Intensifier.

⁴This statement is, strictly speaking, too strong. There are circumstances where the agreement is with an argument in the Ergative or the Dative.

⁵This appears to contradict an earlier remark (following (5a)) about the case of the subject of *d-axar* 'go'. In fact, 'go' is one of a fairly large group of verbs whose subject case varies according to person: A third person subject is Nominative, while first and second person subjects are Ergative.

⁶A screeve is a paradigmatic set differing in person and number, coined from Georgian *mc'k'rivi* by Ak'ak'i Šaniġe. Screeve is less misleading than the more usual 'tense', since paradigmatic sets differ not only in tense, but also in aspect and mood.

⁷Tsova-Tush has a large number of verbs with the formants *-dar* and *-dalar*, among others, which have sometimes been called 'auxiliaries.' Although these suffixes have some properties of auxiliaries (viz., they inflect for person, number, screeve and have special aspectual or modal meanings), they are syntactically different from the auxiliaries defined in this paper. These suffixes function as valence-changing lexical formants.

⁸There is also a Future Series, composed for regular verbs of a Perfective root plus the Present Series suffixes. Since the auxiliaries lack paired perfective roots, none of them have Future Series forms.

⁹This is no doubt due to the fact that *dolo* itself is probably derived from an Imperfective Reported form (suffix *-lo*).

¹⁰Perfectives show a number of grammatical restrictions, the major one being that they do not occur in the Present Series of Screeves.

¹¹With a small number of verbs which are inherently imperfective semantically and lack a paired perfective root (which would be used in forming the future), *latE* is used to express future meaning. For example, *dah*ⁿ *daxa*ⁿ *latE* 'he will be drunk regularly', and *teša*ⁿ *latE* 'he will believe them.'

¹²It is a reasonable guess that the marked order (auxiliary followed by infinitive) in the case of both *latE* and *lati*ⁿ is motivated by features of the discourse. This would be consistent with consultant judgements of unacceptability for the marked order when presented with example sentences without context.

References

- Harris, Alice. 1981. *Georgian syntax*. Cambridge: Cambridge University Press.
- Holisky, Dee Ann. 1985. "A stone's throw from aspect to number in Tsova-Tush." *IJAL* 51: 453-455.
- Kadagije, Davit and Niko Kadagije. 1984. *C'ova-tusur-kartul-rusuli leksik'oni* [Tsova-Tush-Georgian-Russian Dictionary]. Tbilisi: Mecniereba.
- Nichols, Johanna. Forthcoming. "Ingush: A Grammatical Sketch." To appear in *Indigenous Languages of the Caucasus*, ed. by John Greppin et al. Delmar, N.Y: Caravan Press.
- Schiefner, F. Anton von. 1859. "Versuch über die Tusch-Sprache." *Mémoires de l'Académie impériale des sciences de Saint-Petersbourg*, Sixth series, 9:1-160.

The Structure of the Nakh-Daghestanian Verb Root and Verb Stem

Johanna Nichols

0. **Introduction.**¹ There are good studies of verb structure in the Daghestanian languages (Gudava 1959, Magometov 1961), but none for the Nakh branch. This paper is intended to fill that gap, and also to put comparative Nakh-Daghestanian, or Northeast Caucasian (NEC), grammar on a sounder footing in regard to verb structure.

The verb root in NEC languages often consists of a single final consonant, and usually the only element that is shared in cognate sets is that single consonant. This is caused by the fact that nonfinal consonants are often alternating gender-class markers or frozen gender-class markers. These are incontrovertible synchronic and cross-linguistic facts. But sometimes they are used as premises in historical arguments, which has as its result an uncritical approach to etymology in the Nakh family, with nonfinal consonants declared to go back etymologically to class markers at the demand of the investigator rather than at the demand of the language structure. This paper will argue that Nakh nonfinal consonants are not all gender markers, but continue ancient root-initial and medial consonants.

1. **Verb Structure.** Table 1 shows the structure of inflected verbs in the four NEC stocks (causative and similar derivations are not included).²

Table 1. Verb structure in Northeast Caucasian

Nakh	Directional preverb	Spatial preverb	ROOT	—	—	Endings
Avar-Andi- Dido	**	**	ROOT	Determining suffix	Thematic suffix	Endings
Lak	**	**	ROOT	—	—	Endings
Dargi	Spatial preverb	Directional preverb	ROOT	—	—	Endings
Lezghian	Spatial preverb	Directional preverb	ROOT	—	—	Endings

Key: — lacking
** rare

Table 1 shows that preverbs cannot safely be reconstructed for Proto-NEC: preverbs are not found in all branches, and in the families that have them the two preverb types occur in different orders. Furthermore, the preverbs, where present, do not appear to be cognate. The creation of preverbs from postpositions is a transparently recent, and in fact ongoing, development in Nakh. Finally, loss of preverbs without a trace is an unlikely development. For all these reasons, preverbs — and hence preverb slots — cannot be reconstructed in their present form for Proto-NEC. (Evidence will be given in 2.5 for some kind of ancient process involving prefixation, but that process has nothing to do with modern preverbs.)

The ordering of the two preverb types found in the Daghestanian languages is the same as that found in Georgian: spatial or locative preverbs ('in,' 'out,' 'on,' etc.) precede directional or deictic preverbs. The order found in Nakh is shared with Ossetic. Some examples:

Georgian	še - mo - varda in here ran '(s/he) ran in (toward speaker), came running in'
Ossetic	ær - ba - dux + kodtæ here in ran Aux id.
Ingush	ħa - ?a:ra - v-ealar here out v-went 'he came out (toward speaker)'

The two suffix types of the Avar-Andi-Dido branch cannot be reconstructed outside that branch, and clearly represent an innovation there. (Section 2.6 below reviews what evidence Nakh presents for ancient suffixal morphology; it is slim and ambiguous.) The suffixing derivational morphology of the daughter languages makes it plausible to assume that the protolanguage was suffixing, but no derivational suffixes have been reconstructed for Proto-NEC.

The inflectional endings tend to be cognate within families, but very few appear to be cognate across families. Of the six impersonal desinences described as NEC by Troubetzkoy (1929), only two — *-n* and *-r* — have Nakh reflexes; none of the personal desinences do.

The conclusion to be drawn from the facts of Table 1, then, is that for the Proto-NEC verb we can reconstruct with certainty only the root; we can probably reconstruct suffixal position slots for endings and derivational morphology, but almost none of the suffixes themselves.

2. **Root Structure in Nakh.** This section is based on a survey of all the verbs in Maciev 1961, all the verbs in my field and text notes, and a partial search for cognates in Batsbi (using Kadagidze and Kadagidze 1984). It is thus biased toward Chechen; but, since the Nakh family is not deep, this bias is of minimal consequence. Verbs number in the thousands (since causatives, iteratives, etc. are given separate entries in dictionaries), but there are only 260 synchronically irreducible roots. It is these roots that are surveyed here.

2.1. *The verb root canon* for Nakh is:

$$((C_1) V (R)) C_2$$

where C^1 can be an alternating gender marker and C^2 can be *-st* or a geminate.

The variants of this formula and their frequencies are shown in Table 2. (*D* symbolizes the alternating gender marker. There are four markers: *j*, *v*, *d*, *b*; *d* class is the conventional citation form. Examples are from Chechen, with Batsbi cognates, identified by *B.*, adduced where Chechen-Ingush lenition has obscured consonants. The obscured consonants are symbolized by *C.)

Table 2. Root structure variants in Nakh.

Shape	Number	% of 260	Example
Gender-inflected verbs (D = alternating gender prefix)			
DVC	56		D-a:g- 'burn'
DVCC			
-CC = geminate	11		D-yett- 'milk', 'beat'
-CC = -st	3		D-ust- 'measure'
-CC = other	5		D-yelx- 'cry'
total -CC	18		
DVCCC (-rst)	1		D-arst- 'get fat'
D	1		D- 'make, do'
Total	76	29%	

Non-gender verbs, vowel-initial

VC	16	a:l- 'say'
V*C	1	ie ⁿ , B. egar 'get mixed'
VCC:		
-CC = geminate	3	a:xk- 'dig'
-CC = -st	0	
other	3	a:lx- 'comb wool'
total VCC	6	
Total	23	9%

Non-gender, consonant-initial

C	2	g- 'see'
CVC	104	tuox- 'hit'
CV*C	6	lie ⁿ , B. levar
CVCC:		'talk'
-CC = geminate	20	la:tt- 'stand'
-CC = -st	4	q'a:st- 'separate'
other	24	leč'q'- 'hide'
total CVCC	48	
CVCCC (-rst)	1	qarst- 'go, wander'
Total	161	62%
Total non-gender verbs	184	71%

These figures show that over a quarter, but less than a third, of the verb roots show gender prefixation in Nakh. This is the lowest proportion attested in any NEC stock still preserving genders (i.e., to the exclusion of Lezghi, southern Tabassaran, Agul, and Udi). About half of the Andic verbs, and under half of the Avar verbs, distinguish gender (Gudava 1959:17). In Dargi and in the Lezghian languages having genders, grammatical descriptions give the impression that the majority of verbs distinguish gender and those that do not are exceptional; verbs lacking either productive or fossilized gender markers comprise only one of the six structural types distinguished by Magometov 1961 for Tabassaran and Dargi.

Table 2 also shows that monoconsonantal structure is far from typical for the Nakh root. Monoconsonantal roots — those with the structures D, VC, V*C, and C — total only 20, or 8% of the 260 roots. If we add to this the 56 roots with the structure DVC, on the assumption that the gender prefix D is not part of the root and thus DVC roots are monoconsonantal, the total is 76 = 29%.

The verbs in -CC include all those with final geminates. Some of the final geminates are etymologically single consonants: they correspond to Proto-Daghestanian fortis stops and affricates, and to the voiced affricates that yield fortis reflexes in some Daghestanian dialects. Some cognate sets:³

Dagh.	*ʒ	Nakh *tt	in 'stand up, get up':
-------	----	----------	------------------------

Dido	i-z-a	Chechen, Ingush	ʔa:tt-, (h)wott-
------	-------	-----------------	------------------

Lak	i-z-an	Batsbi	qett-
-----	--------	--------	-------

Tabassaran	ʔu-zw-uz		
------------	----------	--	--

etc. (Giginejšvili 1977:84, Bokarev 1981:23)

same correspondence in 'milk (a cow, etc.):'

Avar	c'c'u-jze	Chechen, Ingush	D-yett-
------	-----------	-----------------	---------

Andi	bu-c'c'-in-na	Batsbi	bettar (noun) 'milk cow, sheep, buffalo'
------	---------------	--------	--

Lak	tti-z-in		
-----	----------	--	--

Dargi	dir-z-is		
-------	----------	--	--

Lezghi	acc-un		
--------	--------	--	--

Agul	u-z-as		
------	--------	--	--

etc. (Giginejšvili 1977:84, Bokarev 1981:17)

Dagh.	*ʒ	Nakh *tt	in 'bake, fry'
-------	----	----------	----------------

Avar	be-ž-ize	Chechen, Ingush	D-att-
------	----------	-----------------	--------

Andi	be-ž-idu	Batsbi	D-att-
------	----------	--------	--------

Akhvakh	bi-ž-ur-ut ¹ a		
---------	---------------------------	--	--

Lezghi	čč-ur-un		
--------	----------	--	--

Agul	u-ž-as		
------	--------	--	--

etc. (Giginejšvili 1977:85)

Gemination is a frequent, but not necessary, concomitant of the fortis manner of articulation distinguished among NEC obstruent series. This is true of both Nakh and Daghestanian languages. Thus these sets represent synchronic — and plausibly historical — single fortis consonants, not true geminates. Therefore, at least some of the Nakh final geminates reflect original Proto-NEC single consonants, not clusters. (Many verbs with final geminates, of course, simply lack etymologies.)

If geminate -CC sometimes reflects a single fortis stop, then at least some of the DVCC and VCC roots may also be monoconsonantal. If we add all such roots to the 76 others that are or may be monoconsonantal, the total is 90, or 35% of the total verb root stock. Even by the most

generous possible criteria, then, only about a third of the Nakh verb roots can be regarded as monoconsonantal.

Nakh final *-st* corresponds regularly to Proto-Daghestanian **c*, e.g. in 'measure' (the Batsbi glottalization is a regular secondary development):

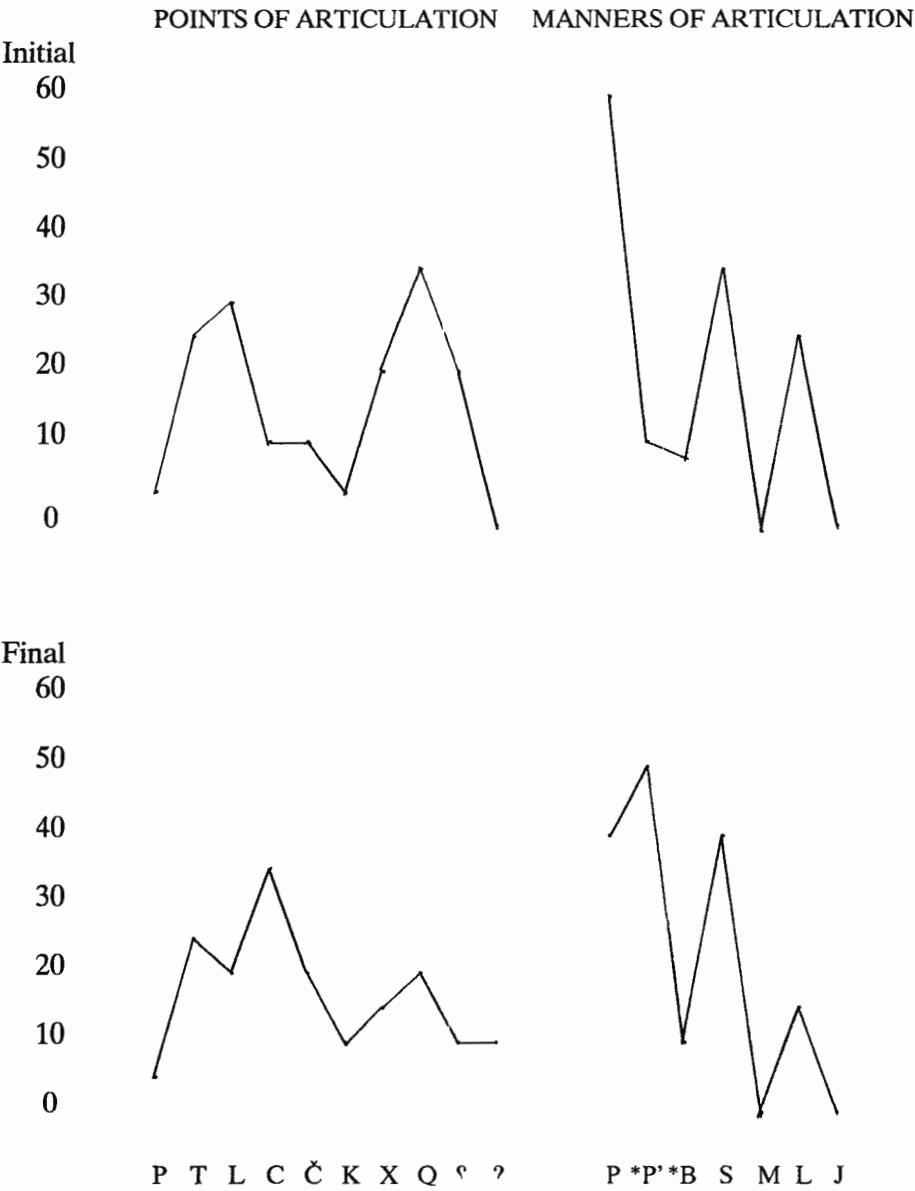
Avar bo-c-ize	Chechen D-usta-	'measure'
Andi ba-s-on-nu	Ingush D-ista-	
Akhvakh ma-č-un-ut ^ʔ a	Batsbi D-ust ^ʔ -	
Lak du-c-in		
Dargi um-c-es		
Lezghi al-c-um-un		
Tabassaran jar-c-uz		
etc. (Giginejšvili 1977:90, Fähnrich 1984:18-19)		

The same correspondence appears in the noun 'bull': Proto-Daghestanian **unc*, Chechen *stu*, Ingush *ust*, Batsbi *p'st'u* (for Daghestanian cognates see Giginejšvili 1977:72, 89, Bokarev 1981:27). An evident fortis analog appears in 'bile': Chechen, Ingush *stim*, Avar *ccin*, etc. (Giginejšvili 1977:71, 125, Bokarev 1981:27). But in other sets Nakh *c* corresponds to Proto-Daghestanian **c*, e.g. 'one' (Daghestanian cognates in Giginejšvili 1977:89; Nakh: Chechen *čhaʔ*, Ingush *caʔ*), 'tooth' (Daghestanian forms from Giginejšvili 1977:70; Proto-Nakh **ca-*). Since there is no evident conditioning environment for a Nakh split of **c* into *c* and *st*, the cluster **st* must be reconstructed for Proto-NEC; in Proto-Daghestanian it merged with **c*, but it is preserved in Nakh. Consequently, at least one -CC verb root, that for 'measure,' can be reconstructed for Proto-NEC.⁴

2.2. *Frequencies and distribution of consonants.* Figures 1 and 2 show the frequencies of initial and final consonants in Nakh verb roots, by point and manner of articulation. Points and manners of articulation are labeled at the bottom with representative consonants. In Fig. 1, laterals (L) and the velar fricative /x/ (X) are treated as separate points of articulation. This was done because both are very frequent and neither is phonetically identical in point of articulation to either of the points adjacent to it: laterals are neither dental (T) nor alveolar (C); and /x/ in Chechen and Ingush is pronounced farther back than /k/ but not as far back as /q/. In Fig. 2, the entry P' symbolizes ejective stops and affricates. These take the form of actual ejectives in initial position. In noninitial position, Proto-Nakh ejectives (other than **q'*, which is preserved intact) are lenited in Chechen-Ingush to their voiced correspondents. The entry P' stands for /q'/ and the voiced reflexes, i.e., it symbolizes the Proto-Nakh ejectives, whatever their reflexes. The top half of each figure

shows the frequency of each point or manner of articulation in root-initial position; the bottom half shows its frequency in final position.

Figure 1. Frequencies of initial and final consonants in Chechen verbs



The following patterns are shown in Figure 1:

Laterals and uvulars are the two most frequent points of articulation in initial position. In final position, the differences are somewhat smoothed out. The alveolars (c, *c', s) are the most frequent finally (although they are infrequent initially), followed by dentals and the postvelars X and Q. Laterals are of only moderate frequency. Labials, true velars (not including /x/), and glottal stops are infrequent in both positions. The only consonants that are notably common in both positions are uvulars.

Voiceless stops, voiceless fricatives, and resonants (primarily /l/) are the dominant initial manners of articulation. In final position, voiceless fricatives and resonants remain frequent, and the reflexes of glottalized consonants are the most common. (It is curious that Proto-Nakh ejectives were most common in root-final position, when their Chechen-Ingush reflexes lose their glottalization in just this position.)

Roots with both initial and final glottalized consonants are rare. I found only the following six (cited in their Chechen-Ingush forms):

- q'a:ž- (*q'a:č'-) 'smile'
- c'i:z- (*c'i:c'-) 'scream, squeal' (onomatopoetic)
- q'ieq'- 'rumble, thunder' (onomatopoetic)
- q'a:d- (*q'a:t'-) 'glance, catch a glance'
- q'ag- (*q'ak'-) 'shine, get decorated'
- q'arz- (*q'arc'-) 'stand out, bug out (eyes); get multicolored (forest)'

Two of the six are onomatopoetic, and all are candidates for categorization as expressive verbs (as that term is defined in Holisky 1981:84, 102). Roots with two ejectives, then, may lie outside the bounds of normal, non-expressive, root structure.

It is interesting that there are no Nakh verb roots with initial sequences of consonant plus pharyngeal. Root structures such as *bʕar*- 'eye,' *čħa*- 'one,' *dʕa*- 'there' are frequent among other parts of speech, but do not occur among verbs.⁵

2.3. Final clusters. A fair number of Nakh roots — a total of 35 — end in clusters other than geminates. The clusters and their frequencies are shown in Table 3.

Table 3. Final clusters.

C ¹	C ²	frequency	remarks
l	x	5	Batsbi -tx
r	t	2	
r	d	2	*rt'
r	c	2	
r	s	2	
r	z	3	*rc'
r	č	2	
r	š	1	
r	ž	2	*rč'
r	ɣ	1	
r	h	1	Batsbi -rl
r	st	1	
total -lC+-rC		24	
s	t	7	note -rst, just above
x	č	2	one cognate to -rc ?
č'	q'	1	
t'	q'	2	
j	?	2	
j	q	1	
n	g	1	
total other		16	

Clusters with a sonorant as first member (-RC) are the great majority. The sonorants are *-r-* and *-l-*, in complementary distribution. There are no clear Daghestanian cognates to such roots. (Although prefinal resonants abound in Daghestanian, especially Lezghian and Dargi, verb roots, they are class markers or fossilized class markers.) Fähnrich 1984:21 compares Chechen-Ingush *talx-* 'spoil' to Avar *tur-* Andi *turd-*, Botlikh *tamd-*, Karati, Akhvakh *tor-*, Bagvali, Tindi *tord-*, Chamali *tan-*. Chechen-Ingush *l* could in principle be cognate to the Avar-Andi-Dido *r* and *n*, since the complementary distribution of Chechen-Ingush sonorants in preconsonantal position could point to an earlier merger.

Chechen-Ingush *x* cannot be cognate to Avar-Andi-Dido *d*; in any event, the latter appears to be a determining suffix. If these roots are cognate, they do not explain the Chechen-Ingush cluster.

The cluster **-st* must go back to Proto-Daghestanian **-st*, as argued above.

There are two cognate sets suggesting that Nakh *-xč* is somehow cognate to, or alternated with, *-rc*: the verbs *marc-* and *čaxč-*, both 'sing, fire' (for the discrepant initials see section 2.5), and a noun set, Chechen-Ingush *a:xča* 'coin(s), money' beside Proto-Daghestanian **arc-* 'silver' (Giginejšvili 1977:26, 68, 90). If there is a systematic connection of *-xč* and *-rc*, then *-xč* is a further representative of the *-RC* cluster type.

The cluster type *-Cq'* (*-č'q'*, *-t'q'*) is interesting. This is the harmonic cluster type common in Kartvelian and Northwest Caucasian. In Nakh it also occurs occasionally in noun roots, but is overall infrequent. One of the verbs in this group, *lač'q'*- 'hide, steal,' may have a cognate in Avar-Andi-Dido **-q'*- 'hide, steal' (Gudava 1959:109, #82). Batsbi *ħač'q'*- 'press, squeeze' (no Chechen-Ingush cognate) may correspond to Avar-Andi-Dido **q'* 'press', etc. (Gudava 110-12, ##83-84). If these are cognates, either the *-č-* is a Nakh addition or it has been lost in Avar-Andi-Dido. Now, Avar-Andi-Dido seems to regularly lose **r* before consonants: cf. Avar. *bac* 'wolf,' Andi *boc'o*, Akhvakh *bač'a*, Dido *boc'i* (and Dargi *bic'/bec'*) beside Chechen-Ingush *borz* (**borc'*), Batsbi *b'arc'* (with adventitious pharyngealization), Lak *barc'* (Daghestanian cognates cited from Giginejšvili 1977:101). Furthermore, the Nakh *-Cq'* clusters are in complementary distribution with the *-RC* clusters: sonorants do not appear before *q'*, and affricates appear only before *q'*. Finally, none of the evidence we have seen so far suggests insertion of prefinal consonants in Nakh. Although this claim is somewhat speculative in the absence of even one cognate set with Daghestanian *-RC*, it is at least possible that the Nakh *-Cq'* clusters also reflect **RC*.

The remaining three clusters are unlikely to be ancient and native. *-jʔ* and *-jq* cannot be reconstructed in that form. No Batsbi cognates to these roots were found, but ordinarily Chechen-Ingush *-VjC* corresponds to Batsbi *-VCiC* (e.g. Chechen *ħajg*, Batsbi *ħabik* 'spoon'). *-ng* is an oddity: this root lacks a Batsbi cognate; and Proto-Daghestanian **-VnC* seems to regularly yield Nakh **-V:C* (see 'stand' and 'bull' in section 2.1).

In summary, if we push the argument based on complementary distribution to its limit, we can see clear Nakh evidence for reconstructing two systematic root-final cluster types in Proto-NEC: **-st* and **-RC*.

The points of articulation of the second consonants of clusters other than *-st* in Nakh are as follows:

dental, alveolar, palatal	18
velar, uvular	11

This is consistent with the overall relative frequencies of these points of articulation in root-final position.

2.4. *Initial clusters.* In contrast to noun roots, Nakh verb roots never begin in clusters.

2.5. *Initial alternations.* Alternations or discrepancies in root-initial consonants are attested in Daghestanian verbs. They always seem to involve gender prefixes or frozen gender prefixes, and they seem to occur between languages rather than within languages. In Nakh, in contrast, they occur within languages, they do not always involve gender prefixes, and in fact a number of them very clearly have no connection with gender prefixes.

I admit a Nakh verb pair as showing an initial alternation only if it meets the following criteria: The meanings are identical or nearly so; the final portions of the roots are identical (except for 'throw', which involves a purely phonological variation); and vowel changes, if present, are those connected with aspect and number ablaut. (One set includes a discrepancy in vowel length.)

The Nakh initial alternations found in this survey are shown in Table 4. (More examples could be found by searching the Chechen data more thoroughly and covering Batsbi and Ingush in the same depth as Chechen.)

Table 4. Initial alternations

D- : q-

Ch. qi - 'realize, perceive, catch on; make it in time; develop, grow,'
* 'arrive'

Ch., Ing. D-ye - 'come' (past, infinitive — suppletive root)

Ch. D-a:st- 'untie, undo'

q'a:st- 'separate (off, out), part' (intrans.)

Ch., Ing. D-añ- 'bring' (present — suppletive root)

qieñ- 'carry, drag' (multiple action)

B. qeñ- 'bring'

Ch. D-wott- 'pour, sprinkle; load, lay (foundation)'

B. D-ott- 'pour'

qott- 'pour, splash (water)'

D- : q- : ʕ- : t- : Ø

Ch. ʕill- 'lie'

D-ill- 'put, lay; spread over; cover' (labile)

D-woll- 'insert' (sg. obj.)

qwill- 'make, create; cast over'

till- 'roof over; name'

Ing. all- 'lie'

qoll- 'put on'

D-ill- (as in Chechen)

Batsbi Doll- (sg. obj.) 'lay, put in, stick in, dig in'

Dill- 'put'

qoll- 'put on (e.g. head), cover, throw on'

oll- 'put on, down, up'

till- 'add (wood to fire); name'

D- : h : Ø-/ʕ-

Ch., Ing. had- 'run' (sg. subj.)

id- 'run; flee, avoid' (multiple action)

D-ad- 'run' (sg. subj.)

Batsbi it'-, ʕit'- 'run'

D- : Ø- : t-

Ch. a:q- 'avenge'

D-ieq- 'pay off, avenge'

taq- 'compensate, pay off, avenge'

D- : t-

Ch., Ing. D-al- 'give' (past, infinitive)

tiel- 'pay'

Batsbi D-ał- 'give'

teł- 'pay'

D- : Ø-

Ch. D-a:t'- 'break off, tear' (intr.)

iet'- id. (multiple action)

Ch. D-ax- 'go, flow'

yex- id. (multiple action)

D- : ʕ-

Ch. ʕitt-, ʕwott- 'stick in, stab'

Ch., Ing. D-yett- 'beat, hit' (multiple action)

D- : g- (?)

Ch., Ing. D-owz- 'know' (*kennen*)

gowza 'skilled, experienced'

Batsbi D-ap'c' 'know' (*kennen*)

gowzan-ča 'specialist, skilled craftsman' (loan from Chechen-Ingush?)

(The meaning of the adjective *gowc'- 'skilled, experienced' may be the missing link between the meanings 'know' and 'preserve, keep' of the Daghestanian cognate set: for the Daghestanian words see Giginejšvili 1977:101.)

l- : γ- : h/θ- : q-

Ch., Ing. la:tt- 'stand' (static)

γa:tt- 'stand up, get up'

Ch. hwott-, Ing. ott- 'assume standing or kneeling position'

Batsbi latt- 'stand' (static)

ott- 'assume standing position'

qett- 'get up' (iterative)

t- : q-

Ch. tas- 'sprinkle, pour'; iter. ti:s- 'throw, hurl, cast'

qoss- 'throw, cast'

qi:s- iterative

qows- multiple action 'rush, pour' (water, rain, etc.)

Ing. tass-

qoss-, qu:s

Batsbi tas- 'throw, fling' (iter. teps-)

qos- 'throw' (iter. qeps-)

l- : θ-

Ch., Ing. iec- 'take, buy'

la:c- 'catch'

An example of a less clear case, plausible but not included in the tabulation below because of discrepancies in meaning and vowel length, is:

D- : h-

Ch. D-a:g- 'burn; be afraid, beware'

ħag- 'wish strongly, envy; be thirsty'

Certain consonants recur often in these alternations. Frequencies of consonants in alternations are shown in Table 5.

Table 5. Frequencies of consonants in initial alternations.

consonant	sets	alternation types
D	12	8
q	7	5
t	4	4
∅	4	4
l	2	2
ʕ	2	2
ħ	2	2
ʕ/∅	1	1 *
∅/h	1	1 *
h	1	1
ɣ	1	1
g	1	1

* These two initials may themselves be alternations.

The fact that initial alternations occur at all leads one to suspect ancient prefixation, now frozen. That some consonants are conspicuously frequent among these alternations supports this suspicion: these are the likely candidates for ancient prefixes. It may be, then, that at least some instances of Nakh CVC roots go back to *C-VC or *CV-C.

Of these alternating consonants, *D* is known to be a prefix: it is the gender marker (itself an alternation of *j*, *ʋ*, *b*, *d*). It is found in alternation sets with every other alternating consonant, and it is the most frequent alternating consonant. It can be concluded that wherever there is evidence for ancient prefixation, gender marking gets into the act.

It is not surprising that *q*- is also frequent: as Figure 1 showed, *q*- is generally of high frequency. (Recall that Figure 1 lumped *q*- with *q*'- and

γ - under the point of articulation Q; q - is not as frequent as the more generic Q.) On the other hand, l - is more frequent overall than q -; yet l - is distinctly less frequent than q - among the initial alternations. Thus the involvement of q - in initial alternations is not an automatic result of phonological structure. There may have been a q - prefix.

Equally likely, though, is a scenario where gender inflection arose not by prefixation of gender markers onto vowel-initial roots (the traditional synchronic analysis, projected diachronically) but by *replacement* of etymological initial consonants with gender markers. On that analysis, q - is the consonant most prone to replacement by gender inflection. This account explains the ranking of consonants, with D at the top, shown in Table 5. It assumes that gender inflection was a secondary addition to verb roots, and that it invaded the lexicon gradually, favored in certain phonological environments.

But this account does not explain the initial alternations involving more than one consonant in addition to D-, nor does it explain the alternations such as t - : q - where there is no gender inflection. Such alternations still point to ancient prefixation, independent of (and perhaps prior to) gender inflection. They indicate that gender alternations are the inflectional manifestation of a tendency that also existed in derivation. The spread of gender inflection through the lexicon would obviously have been favored by pre-existent root-initial consonant alternations.

Support for initial alternations which involved prefixation and were independent of gender inflection comes from Bouda 1949:15ff., where a set of Lak roots with possibly prefixal initial l - is cited. Bouda sees the l - as preverbal partly on etymological grounds and partly because it precedes the gender marker in the inflected verb. I cannot evaluate the etymological argument; the argument based on position could be falsified if it could be proven that the root-medial gender marker was never a prefix, but results from the replacement of root-medial resonants by gender markers. Consistent with the account to be given below, I assume that the l - is prefixal rather than preverbal. It is significant that one of Bouda's l - verbs is cognate to a Chechen-Ingush set with an initial alternation: Lak *lasun* 'take, buy', Chechen-Ingush *la:c*- 'catch, grab' and *iec*- 'take, buy'. Another is cognate to a Chechen-Ingush verb with initial gender prefix: Lak *laxxan* 'put on', Chechen-Ingush *D-u:x*- id; 'dress'. Even more significantly, a third displays a Lak-internal initial alternation and has a gender-prefixed Chechen-Ingush cognate: Lak *lišin* 'wipe, wipe off' : B-*ašin* 'sweep (floor, etc.)' (where B- symbolizes the gender inflection) : Chechen-Ingush *D-a:š*- 'shave'.⁶ Bouda also finds six verbs with possibly prefixal tt -. The regular Nakh correspondent to this

consonant would be *d-*, which coincides with the gender prefix for the residual or unmarked gender category and would surely have been reinterpreted as a gender prefix.⁷ Significantly, one of the Lak verbs with prefixal *tt-* has an apparent cognate to a Chechen-Ingush gender prefixed verb: Lak *ttiḡin* ‘suck(le)’, Chechen-Ingush *D-aq-* id., another has a cognate with no initial consonant: Lak *ttizin* ‘milk (a cow, etc.)’ : Chechen-Ingush *uoz-* ‘milk; pull’ (Giginejšvili 1977:84; see also Schulze 1985).

If *q-*, *l-*, *t-*, etc. are ancient prefixes, there is no evidence in the verbs of Table 4 for what those prefixes may once have meant. It may be significant that most (although not all) of the verbs have to do with motion and stance, that some are paired for transitivity, and that gender inflection seems to be favored by transitive roots and disfavored by multiple aspect.

2.6. *Final alternations.* A few verb root sets show alternations of final consonants. These are fewer in number than the sets with initial alternations. They always involve close phonological relations — typically, a difference of one or two distinctive features — and hence are likely to reflect phonological changes rather than suffixation. The sets are shown in Table 6. (Two of them also involve initial alternations; these were not included in Table 4.)

Table 6. Final alternations

-c : *-čʰ

Ch. *Da:ž-*, *Da:c-* both ‘graze’

*-tʰ : -l

Ch. *xa:d-* ‘break (off), tear (off)’

xa:l- id. (in certain contexts)

-x : -q

Ch. *ħaq-* ‘rub, spread (on); saw’ (back-and-forth motion)

ħax- ‘spread on, smear, plaster’

-tt : *-cʰ (?) (i.e. Proto-NEC *z : *cʰ; Proto-NEC alternations of this type at other points of articulation are mentioned by Giginejšvili 1977:135)

Ch., Ing. *Dett-* ‘milk (cow)’; Batsbi *bettar* ‘milk cow’

Ch, Ing. *uoz-* ‘pull, weigh; milk’; Batsbi *ocʰ-* (see 2.5)

-rc : -xč

Ch. marc- 'singe'

čaxč- 'get singed' (see 2.3)

It should be noted that infixation of multiple and iterative markers (*Y, *W), a regular process, causes devoicing or deglottalization of the final consonant. This is recent and apparently automatic, and does not count as a final alternation of the type discussed here. Examples from Chechen:

qwoss- 'throw'

iterative qi:s-

multiple qows- (B. qeps-)

Diž- 'lie down' (sg.)

Di:š- id. (pl.)

The following words are the only set with an alternation possibly involving a shift of position. Since the phenomenon is unique, the similarity is probably accidental.

Ch, I a:l- 'say' le*W 'talk'

Batsbi ał lev-

In summary, final alternations are rare; the sets are not as convincing as those for initial alternations; the alternations involve minimal phonological differences, and hence appear purely phonological rather than morphological. They do not suggest Proto-NEC suffixation.

2.7. *Gender markers.* The gender prefixes in Nakh are always transparent and regular. They show no allomorphic variation and virtually no allophonic variation. They are almost never frozen; the only example of possibly frozen gender markers is Ch. *ba:x-*, Ing. *ja:x-* 'talk' (*b-* and *j-* are two of the gender markers, but in these verbs they do not agree and do not alternate). Gender inflection is certainly not dying out, and may even be productive and spreading in Nakh. It is found in initial position only; in Nakh there is no evidence that root-final consonants ever alternated.

3. *Comparative Analysis and Reconstruction.* The synchronic Nakh verb root canon, and the canon that can be reconstructed for Proto-Nakh and even pre-Proto-Nakh, is $*(C_1 V R) C_2$, where $*C_2$ includes geminates and the cluster **st*. The near absence of evidence for $*C_1$ and $*R$ in the Daghestanian languages can be explained by assuming that gender markers regularly replaced both $*C_1$ and $*R$ in at least Lezghian and Dargi, and that $*R$ was subject to phonological attrition (especially in

Avar-Andi-Dido), while Nakh gender markers replaced $*C_1$ in some verbs but never replaced $*R$. Many instances of Proto-NEC $*R = *n$ could be lurking behind Nakh verbs with long root vowels; as mentioned above, $*VNC$ apparently yields Nakh $V:C$.

Gender inflection replaces more consonants, affects more verbs, and is phonologically less transparent, in the easternmost languages (the Lezghian and Dargi families) than in Nakh. This distribution suggests that gender agreement in verbs spread through NEC from east to west, affecting Nakh significantly later than it affected Lezghian. The evidence for this spread, taken together with the verb canon reconstructable for Nakh, calls into question the traditional assumption that gender agreement in verbs is Proto-NEC. (It says nothing about whether nouns had gender; it merely questions the antiquity of the modern formal category of verbal agreement.) It suggests a Proto-NEC root canon of $*(C_1 V R) C_2$ or even (since 'vowel'-initial roots actually begin with a phonemic glottal stop in some languages, including Nakh) $*C_1 V (R) C_2$, with $*C_1$ and $*R$ almost entirely obliterated by gender inflection and phonological attrition in the easternmost languages.

At least for Proto-Nakh, we can reconstruct some kind of prefixation that was distinct from gender inflection. The other stocks are unlikely to provide firm evidence for this other kind of prefixation, just because gender inflection was more pervasive in its replacement of etymological root initials. On the other hand, the very fact of replacement and its pervasiveness may be indirect evidence for ancient prefixation of other types: as was mentioned above for Nakh, the existence of prior root-initial alternations would have favored the rise and spread of root-initial gender-marking alternations.

Most work on NEC history and reconstruction has been based primarily on Daghestanian data, and the Nakh correspondences have been poorly understood. This has distorted our view of Proto-NEC, and it has distorted it in a direction maximally inhibiting to reconstruction, since the facts reviewed above show that the Nakh branch is the most conservative of all the NEC branches in its root structure.

4. Implications for Subgrouping. Verb structure provides no evidence for subgrouping within the NEC family: the sharings and discrepancies surveyed above follow from geographic distribution and sometimes crosscut families, and many of them can be explained by a post-Proto-NEC spread of gender inflection from east to west across the already differentiated daughter stocks. On the other hand, the $*st$ cluster reconstructed above — a cluster which draws its claim to reconstructability from the fact, established here, that the Proto-NEC verb was *not* regularly monoconsonantal — does give evidence for

subgrouping. Its retention in Nakh, as against its merger with *c elsewhere, supports the traditional first split of Proto-NEC into Nakh vs. Daghestanian.

University of California, Berkeley

Notes

¹Some of the research for this project was done in Tbilisi, Georgia in 1984 as a participant in the Exchange of Senior Scholars between ACLS and the Soviet Academy of Sciences, with the support of the International Research and Exchanges Board. I am grateful to the Oriental Institute of the Georgian Academy of Sciences and its director Thomas V. Gamkrelidze for making the research possible, and to the Linguistics Institute of the Academy for consultations and library resources. The remainder of the research was made possible by sabbatical leave from the University of California, Berkeley.

²Sources for entries: Nakh: my own field work and literature survey. Avar-Andi-Dido: Gudava 1959, Gamzatov 1983. Lak: Victor Friedman, personal communication. Dargi, Lezghian: Magometov 1961, 1983.

³Daghestanian forms are cited from Giginejšvili 1977, with some retranscription (e.g. fortis consonants written as geminates, lateral affricate written tʃ). The hyphens in the Nakh forms mark morpheme boundaries. The hyphens in Giginejšvili's Daghestanian forms sometimes coincide with morpheme boundaries, but sometimes they split morphemes and should therefore be regarded as segmenting out the root consonant at issue (W. Schulze, personal communication). Hyphens in Nakh forms mark morpheme boundaries.

⁴Besides 'measure' there are six other verbs in -st and two in -rst. I will list them all in the hope of spurring searches for their cognates: Ch-I. *Da:st*- 'untie, undo', *Dyest*- (= *Dest-) 'swell up, rise (water)', *q'a:st*- 'separate; be distinguished, defined', *last*- 'rock, shudder', *list*- 'reel (up, off), wind' (thread), *ħast*- 'caress', *Darst*- 'get fat, gain weight', *qarst*- 'go, wander'.

⁵The pharyngeal in such roots shares the manner of articulation of the initial consonant: /h/ follows voiceless (aspirated) obstruents, /ʕ/ follows voiced and glottalized consonants. Phonologically, such roots can be analyzed either as having initial sequences of consonant plus pharyngeal or as having pharyngealized vowels. Phonetically, there is some variation: in Ingush they seem always to be pronounced with pharyngealized vowels, while in Chechen and Batsbi the voiced pharyngeal is often audible as a glide following the consonant and the voiceless pharyngeal (as in Ingush) simply amounts to pharyngealization of the aspirated release.

⁶The first two sets are mentioned by Bouda. The first displays an irregular correspondence (Lak points to *s, Nakh to *c) which falls into an attested Proto-NEC alternation type: see Giginejšvili 1977:133, 135.

⁷There is only one verb in Chechen with a non-alternating, fixed initial *d*-. *dieb*- 'multiply [of animals]', in which the *d*- is surely a frozen gender prefix agreeing with the

word meaning 'young of animals'. The absence of any clear instances of root-initial *d*-, as opposed to *D*-, is in itself good evidence that root-initial consonants were replaced by gender markers.

References

- Bokarev, E. A. 1981. *Sravnitel'no-istoričeskaja fonetika vostočnokavkazskix jazykov*. Moscow: Nauka.
- Bouda, Karl. 1949. *Lakkische Studien*. Heidelberg: Winter.
- Fähnrich, Heinz. 1984. 'Zur Stellung des Nachischen in der nachisch-daghestanischen Sprachfamilie.' *Sprachen Kaukasiens*, 16-21. (Friedrich-Schiller-Universität Jena, Wissenschaftliche Beiträge.)
- Gamzatov, R. E. 1983. "O preverbno-posleložnoj sisteme mazadinskogo (taškogo) govora avarskogo jazyka." In Tabulova and Temirova 1983, 236-42.
- Gigineišvili, B. K. 1977. *Sravnitel'naja fonetika dagestanskix jazykov*. Tbilisi: University Press.
- Gudava, T. E. 1959. *Sravnitel'nyj analiz glagol'nyx osnov v avarskom i andijskix jazykax*. Maxačkala: Dagestanskij filial AN SSSR.
- Holisky, Dee Ann. 1981. *Aspect and Georgian Medial Verbs*. Delmar, N.Y.: Caravan Books.
- Kadagidze, David and Niko. 1984. *Cova-tušinsko-gruzinsko-russkij slovar'*. Tbilisi: Mecniereba.
- Maciev, A. G. 1961. *Čečensko-russkij slovar'*. Moscow: Gos. izd. inostrannyx i nacional'nyx slovarej.
- Magometov, A. A. 1961. 'O stroe glagola v tabasaranskom jazyke.' *Voprosy izučenija iberijsko-kavkazskix jazykov*, 227-39. Moscow: AN.
- _____. 1983. "Sistema posleložnyx padežej i preverbov v darginskom jazyke (po dialektnym dannym)." In Tabulova and Temirova 1983, 191-202.
- Mejlanova, U. A. 1984. *Buduxsko-russkij slovar'*. Moscow: Nauka.
- Schulze, Wolfgang. 1985. "Zu den Bezeichnungen für 'Kuh' in den ostkaukasischen Sprachen." H. M. Ölberg and G. Schmidt, eds., *Sprachwissenschaftliche Forschungen: Festschrift für Johann Knobloch*, 437-42. (Innsbrucker Beiträge zur Kulturwissenschaft, 23.) Innsbruck.
- Tabulova, N. T. and R. X. Temirova, eds. 1983. *Sistema preverbov i poslelogov v iberijsko-kavkazskix jazykax*. (Materialy 8-oj regional'noj sessii po issledovaniju sistemy i istorii iberijsko-kavkazskix jazykov.) Čerkessk: Karačaevo-čerkeksskij NII istorii, filologii i ekonomii.
- Troubetzkoy, N. 1929. "Notes sur les désinences du verbe dans les langues tchéchénolesghiennes (caucasiennes-orientales)." *BLSP* 29:3.153-71.

The Subjunctive in Classical Armenian: Significant Differences Between Eznik and Elišē

Alfred G. Paludis

Eznik's *Etc alandoc'* is considered to be in essence the first original work in Armenian. Although it relies heavily on foreign sources, the text seems to have a pureness of style and vocabulary which sets it apart from earlier and later writings. In collecting some basic statistics on the use of the subjunctive in classical Armenian, it was discovered, however, that some disparities exist in usage between Eznik and the biblical texts he quotes. The most striking example is the absence of aorist subjunctive forms in the conditional vs. twenty three in quoted material. When these statistics were compared to those compiled from Elišē's *Vasn Vardanay ew Hayoc' paterazmin*, it became obvious that such disparities occur not only between Eznik and his sources, but also with other classical texts.

No consistent statistical work has been done throughout classical Armenian, so any study of aspect and mood may be considered somewhat impressionistic and less authoritative than might be wished. We think it important to offer a potential framework for verbal analysis in Armenian so as to heighten the awareness of the extent of data upon which subsequent descriptions should be based. The following guidelines have been found useful in defining the various categories:

- A. **Independent:** those clauses having no overt subordinating element, other than *(e)t'e* 'that' after verba diciendi et sentiendi;
 - 1. **Interrogative:** clauses with *paroyk* '?' on one or more words (negation *oč'*);
 - a. **Pronominal:** *paroyk* rests on *o(r)* 'qui(s)', *zi(nč')* 'quid', or one of their declined forms, or a phrase containing them;
 - b. **Adverbial:** *paroyk* rests on an interrogative adverbial, e.g., *usti* 'whence', *yoržam* 'when';
 - 2. **Noninterrogative:** all others, with subclasses distinguished by use of either *mi* 'ne' or *oč'* 'non';
- B. **Subordinate:** clauses introduced by a word or phrase implying dependence on the previous clause;

1. **Relative:** clauses introduced by a relative pronoun or adjective or their compounds, e.g., *or ok* 'whoever', with subtypes defined by the use of the negative, as in A2 above.
2. **Conjunctive:** clauses introduced by one of several types;
 - a. **Conditional:** apodoses introduced by *(e)t'e* 'si' (neg. *oč'*);
 - b. **Adverbial:** clauses introduced by adverbials, e.g., *usti* 'whence', *yoržam* 'when' (neg. *oč'*);
 - c. **zi** 'so that': final, result or consecutive clauses, with subtypes defined by the negation, as in A2;
 - d. **X+zi:** clauses introduced by an element followed by *zi* as a subordinator, e.g., *vasn zi* 'because';
 - e. **(e)t'e** 'that': clauses which are neither conditional nor dependent on verba dicendi et sentiendi (neg. *oč'*);
 - f. **X+(e)t'e:** clauses introduced by an element followed by *(e)t'e* as a subordinator, e.g., *orpēs t'e* 'as if' (neg. *oč'*).

The statistical compilations were made using Eznik (1959) and Elišē (1968). In arranging the data, all material from biblically identifiable quotation has been disregarded in the computations, although it is included in the tables, enclosed in parentheses. Decimals express the percentage of examples in each category based on (1) all examples of that tense in a given work and (2) all examples of the immediate supercategory. In computing increase or decrease Eznik is assumed to be the standard.

The most astounding fact to arise from the data is the reversal in attestation. In Eznik, the aorist accounts for only one fifth of all examples, whereas in Elišē, it comprises nearly two thirds. An important subsidiary trend is the large increase in usage of the present in independent, specifically noninterrogative, clauses in Elišē. A slight increase of the aorist in interrogative clauses is also evident in the latter. These discrepancies might be conditioned by the kind of texts which form the basis of this investigation. Eznik is highly polemical, derisive, and perhaps atypical in comparison to Elišē's descriptive history. It remains to be seen whether these statistics have any grammatical weight or are only stylistically significant.

The results for the conjunctive category are surprising. Subordinate clauses in two contemporary texts of some length may be expected to have less deviating patterns of usage. As can be seen in Table 1, Eznik does not use the aorist in conditional clauses, except for quoted material, so an increase here is not extraordinary; however, the fact that this is accompanied by a sharp drop in the two categories X+*zi* and *(e)t'e* (B2d, e), rather than a general decline in all others, is perhaps indicative of

changes taking place in the grammatical/syntactic framework. Even more interesting is the large decrease of the present in adverbial clauses (B2b) and the gain in *X+(e)t'e* clauses (B2f). It is known that the latter category expanded considerably in Cilician Armenian at the expense of all other nonconditional categories. Since *zi* clauses (B2c) still seem to be quite stable, these statistics may document the beginnings of the drift toward Middle Armenian structure.

With the limited amount of data at our disposal, it would be difficult to make any consistent statements about the historical development of Armenian. In general, *Elišē* gives the impression of being more similar to the Armenian Bible than Eznik. This is unusual, if the latter is indeed the earlier text. A slight drift toward the structure of Middle Armenian in *Elišē* would be consistent with its somewhat later date. Further investigation utilizing the Bible and other literary works may help to clarify this problematic relation, since these differences do appear worthy of explanation and significant, especially in the historical context of Armenian.

University of Chicago

References

- Eznik de Kolb. 1959. *De Deo*, L. Maries et Ch. Mercier, eds. (Patrologia Orientalis, v. 28, f. 3-4). Paris: Firmin-Didot.
- Elišē. 1968. *Vasn Vardanay ew Hayoc' paterazmin*. 2nd ed. Jerusalem: Dproc' Surb Yakobeanç'.

CATEGORY	EZNIK			ELIŠĖ			CHANGE
	no.	%ao.	%cat	no.	%ao.	%cat	%
A. Independent	92(19)	47.1	47.1	180(14)	44.7	44.7	2.4-
1. Interrogative	4	2.0	4.3	23 (3)	5.7	12.7	8.4+
a. pronominal	--	--		12	2.9		
b. adverbial	1	0.5		9	2.2		
2. Noninterrog	88(19)	45.1	95.6	157(11)	39.0	87.2	8.4-
<i>mi</i>	10	5.1		22	5.4		
<i>oč'</i>	1	0.5		6	1.4		
B. Subordinate	103(46)	52.8	52.8	222 (7)	55.2	55.2	2.4+
1. Relative	10 (8)	5.1	9.7	15 (1)	3.7	6.7	3.0-
<i>mi</i>	1	0.5		2	0.4		
<i>oč'</i>	1	0.5		--	--		
2. Conjunctive	93(38)	47.6	90.2	207 (6)	51.4	93.2	3.0+
a. conditional	--(23)	--	--	28	6.9	13.5	13.5+
b. adverbial	7 (3)	3.5	7.5	23	5.7	11.1	3.6+
<i>oč'</i>	2	1.0		--	--		
c. ZI	59(12)	30.2	63.4	132 (5)	32.8	63.7	0.3+
<i>mi</i>	4	2.0		26	6.4		
<i>oč'</i>	--	--		1	0.2		
d. X+ZI	12	6.1	12.9	7	1.7	3.3	9.6-
e. (E)T'E	14	7.1	15.0	12 (1)	2.9	5.7	10.3-
<i>oč'</i>	--	--		3	0.7		
f. X+(E)T'E	1	0.5	1.0	5	1.2	2.4	1.4+
<i>oč'</i>	--	--		2	0.4		
TOTAL	195(65)	100.0	19.7	402(21)	100.0	65.9	46.2+

Table 1. Aorist Subjunctive: Statistics

CATEGORY	EZNIK			EĻIŠĒ			CHANGE
	no.	%ao.	%cat	no.	%ao	%cat	%
A. Independent	138(12)	17.4	17.4	80	38.4	38.4	21.0+
1. Interrogative	124	15.6	89.8	50	24.0	62.5	27.3-
a. pronominal	59	7.4		18	8.6		
b. adverbial	47	5.9		13	6.2		
2. Noninterrog	14(12)	1.7	10.1	30	14.4	37.5	27.3+
<i>mi</i>	--	--		5	2.4		
<i>oč'</i>	1	0.1		1	0.4		
B. Subordinate	655(43)	82.5	82.5	128	61.5	61.5	21.0-
1. Relative	116 (9)	14.6	17.7	24	11.5	18.7	1.0+
<i>mi</i>	--	--		2	0.9		
<i>oč'</i>	17	2.1		--	--		
2. Conjunctive	539(34)	67.9	82.2	104	50.0	81.2	1.0-
a. conditional	143(10)	18.0	26.5	27	12.9	25.9	0.6-
b. adverbial	139 (2)	17.5	25.7	10	4.8	9.6	16.1-
<i>oč'</i>	19	2.3		1	0.4		
c. ZI	102	12.8	18.9	22	10.5	21.1	2.2+
<i>mi</i>	12	1.5		5	2.4		
<i>oč'</i>	4	0.5		1	0.4		
d. X+ZI	19	2.3	3.5	3	1.4	2.8	0.7-
e. (E)T'E	95	11.9	17.6	22	10.5	21.1	3.5+
<i>oč'</i>	10	1.2		--	--		
f. X+(E)T'E	41(22)	5.1	7.6	20	9.6	19.2	11.6+
<i>oč'</i>	5	0.6		3	1.4		
TOTAL	793(55)	100.0	80.2	208	100.0	34.0	46.2-

Table 2. Present Subjunctive: Statistics

Class Inflection and Related Categories in the Caucasus⁰

Karl Horst Schmidt

§1. As is generally known, typological contrasts exist between the East Caucasian (EC) languages, with noun classes, impersonal verb inflection, and the use of case and postpositions, and the West Caucasian (WC) languages, which have polypersonal verb inflection and express syntactic relations by means of preverbs. The South Caucasian (SC) or Kartvelian languages occupy an intermediate position. They have polypersonal verb inflection and preverbs in common with the WC languages, while sharing the use of cases and postpositions with the EC languages (Deeters 1957).

§2. Most important among the features of EC is perhaps the selective differentiation of substantives as to noun classes, a typological peculiarity which is shared by the Bantu languages of Central and South Africa, but not by the WC and SC languages: "die Kondordanz zwischen Regens und Rektum, Subjekt und Prädikat geschieht dadurch, daß Attribute und Prädikate durch lautlich verschiedene Klassenzeichen (KZ) auf die Klasse der übergeordneten Satzteile Bezug nehmen" (Deeters 1963:46). The process of classification is dominated by the substantive as the selective category (Hockett 1958:230), which itself is not characterized by a class marker (CM). Xajdakov 1980:4 stresses its *skrytyj xarakter*. Class as an *inflectional category* (Hockett 1958:230) is restricted to the parts of the sentence which depend on the substantive (in particular to the attributes) and to the verbs connected with it.

- (1) Chechen *cʰa stag xilla nowqa w-öduš* 'a man (*stag*) was walking (*w-öduš*) on the road' vs. *b-jeana qxunna düʰal cʰa läʰa b-älla* 'having crawled to meet (*b-jeana*) him a snake (*läʰa*) appeared (*b-älla*)' (Jakovlev 1940:308)

As the Chechen examples of (1) show, the substantives *stag* 'man' and *läʰa* 'snake' are unmarked themselves, but their different classification is expressed by the CM *w* or *b*, respectively, attached to the relevant verbs.

On the other hand, there seem to be exceptions to the rule that substantives show no class inflection.

(A) There are sets of nouns in which the distinction between male/female and human/animal is marked merely by the initial consonant, which formally corresponds to the CM *w*, *j* or *b* of classes I, II or III (cf. (6) below).

- (2) Avar *was* 'son, boy' : *jas* 'daughter, girl' : *bas* 'cub'; cf. *was w-ačuna* 'the boy comes', *jas j-ačuna* 'the girl comes'; *wac* 'brother' : *jac* 'sister' : *bac* : 'brother, sister (of an animal)'; Chechen *waša* 'brother' : *jiša* 'sister'; *wo*^o 'son' : *jo*^o 'daughter'; Bats *wašo* 'brother' : *jašo* 'sister'; *woh* 'son' : *joħ* 'daughter'; Akhvakh *waša* 'son' : *jaše* 'daughter' : *baša* 'cub'; *wac:u* 'brother' : *jac:u* : *bac:u* etc. (Cf. Čikobava/Cercvadze 1962:240; Klimov/Alekseev 1980:274f.; Timaev 1983:107f.)

In Dargwa, moreover, the same type of differentiation is applied to parts of the body (class II being marked by CM *r*):

- (3) *wah* 'face (of a man)' : *rah* 'face (of a woman)' : *bah* 'face (of an animal)'; *way* 'waist (of a man)' : *ray* 'waist (of a woman)' : *bay* 'waist, middle' (Klimov/Alekseev 1980:274f.; Xajdakov 1980:202f. As to the names for parts of the body in Dargwa cf. Magometov 1982:131).

In addition to this Xajdakov (1980:10f.) distinguishes two chronological periods in Dagestan class inflection:

- (1) "Klassnyj ètap, kotoryj xarakterizovalsja tem, što grammatičeskoe čislo vyražalos' preimuščestvenno s pomošč'ju samyx klassnyx èksponentov, lokalizovannyx v imenax suščestvitel'nyx" (p. 10); (2) "Perexod ot sintetičeskogo sposoba vyraženijsja leksičeskix klassov k sintagmatičeskomu ... ètot process s očevidnost'ju byl svjazan s utratoj KĖ (=klassnyx èksponentov) v substantivax s odnovremennym vozrastaniem roli kategorii morfologičeskogo klassa v soglasovatel'noj sisteme.... Èto odnovremeno označalo perexod ot polifunkcional'noj klassno-čislovoj kategorii ... k ordinarnoj čisto čislovoj kategorii..." (p. 11).

On the other hand, old class inflection in substantives can hardly be proved by the evidence of (2) and (3) above; Klimov/Alekseev 1980:275 mention the possibility that some of the antonymous pairs (*brother/sister*, *son/daughter*) may be of verbal origin, a hypothesis which does not seem very probable as long as we lack the evidence of the relevant verbal roots. Instead, one might rather think of these types as alliterating reproductions following the model of adjectival class inflection. From a typological point of view, the differentiation as to sex in (2) recalls the so-called *Movierung* in IE languages: Greek *hetaĩros* : *hetaĩra*, Latin *victor*

: *victrix* etc. (cf. Wackernagel 1928:10f). A later transfer of the CM from man or animal to the corresponding parts of the body might well explain type 3 (in (3)).

The evidence of type 2 *was/jas/bas* — agrees neither with an inflection paradigm, as in Chechen *w-öduš*, *b-jeana*, *b-älla* in (1), or Avar *w-ačuna*, *j-ačuna* in (2), nor with a *derivational category* (Kuryłowicz 1964:33): in contrast to Latin *vic-tor*, *vic-trix* which are derived from the root *vīc-*, *vic-* (*vinco*, *vīcī* etc.) we lack the evidence of a root **-as* as the basis for *was/jas/bas*.

(B) Evidence of allomorphs of the ergative determined by semantic criteria (*masculine rational* vs. *feminine rational* vs. all other categories) is found in a few EC languages (Avar, Lezgian, Archi):

- (4) Ergative: Avar *wac:-as*: 'brother' : *jac:at* : 'sister' : *ci-c:a* 'bear'; Archi *uš-mu* 'brother', *baba-mu* 'aunt' : *doš-mi* 'sister', *ikw-mi* 'heart' (Čikobava/Cercvaze 1962:110ff.; Klimov/Alekseev 1980:14f.; Xajdakov 1967:611; 1980:135; Kibrik et al. 1977, II:18ff., 59; Xaxaze 1979:114).

These variants of the ergative case cannot be very old, as the CM of the predicative syntagm as a rule does not agree with the ergative, but with the unmarked *absolute/indefinite* case:

- (5) Tabasaran *izu d-isnu-za bay* 'I caught the *lad*' vs. *izu b-isnu-za žaqa* 'I caught the *bird*'; *ermi ti-r-xnuw* 'the *man* flew' vs. *žaqa ti-w-xnuw* 'the *bird* flew' (Magometov 1965:198; Schmidt 1972:451f.; Xajdakov 1980:4f.).

The number of classes ranges from zero (Lezgian, Agul, Udi) (Xajdakov 1980:5) to ten (Akki dialect of Chechen) (Timaev 1983:74). North Tabasaran distinguishes two classes (*human* vs. *non-human*) (Magometov 1965:80); Avar has three classes, marked by the CM I *w*, II *j*, III *b*, and a neutralized plural in *r*, while four-class systems are found especially in Dagestan, e.g., in the singular of six languages of the Lezgian group: Archi, Tsakhur, Rutul, Khinalug, Kryzy, Budukh (Xaxaze 1984:5). Furthermore, Deeters (1963:46) postulated four classes for the Proto-EC reconstruction model:

(6)		Singular	Plural
I.	Masculine rational	w	b
II.	Feminine rational	j	b
III.	Other individuals (animals, plants, material things)	b	d
IV.	Material nouns, collectives	d	d

In the Veinakh languages the number of classes is increased by additional semantic principles and new combinations of CM: we find eight classes in Bats, while both Chechen and Ingush have six. But some of the classes are based on very scant evidence. To cite just two examples given by Timaev (1983:47), in the Akki dialect of Chechen, class VII is only represented by a single word, *ber* 'child', while *ajgar* 'stallion' can be included both in class VIII and in class IX.

A number of hypotheses on the origin and development of class inflection are presently being discussed:

(1) Secondary differentiation of singular and plural (Xajdakov 1980:55; Timaev 1983:187): in Andi the old class inflection with the same morphemes in singular and plural is supposed to have been largely preserved; it is applicable to the three classes of Lower Andi (I *w*, II *j*, III *b*), while in most of the dialects of Upper Andi we observe a five-class system with one class, III, differentiated as to number:

- (7) I *w-očuxa ima* 'the big father' : Pl. *w-očuxol imobil*
 II *j-ečuxa ila* 'the big mother' : Pl. *j-ečuxol ilobil*
 III *b-ečuxa kotu* 'the big horse' : Pl. *j-ečuxol kotil*
 IV *b-ečuxa hinčo* 'the big stone' : Pl. *b-ečuxol hinčobil*
 V *r-ečuxa haq:u* 'the big house' : Pl. *r-ečuxol haq:obil*

(Cercvaze 1965:94; 1967:280).

(2) Gradual development of class inflection: early differentiation as to the dichotomies *animate* vs. *inanimate*, *human* vs. *non-human*, *masculine* vs. *feminine* (II): the Botlikh class system is structured in the singular according to *human* (I + II) vs. *non-human*, in the plural according to *animate* vs. *inanimate*:

- (8) Sg. I *w* : II *j* : III *b*; Pl. *r* (*animate*) : *b* (*inanimate*) (Gudava 1967:295; Xajdakov 1980:56f.).

(3) Development of the grammatical category of *person* in some of the EC languages: the Proto-EC verb is characterized by impersonality combined with class inflection. Therefore the development of a finite verb by affixing personal markers must be regarded as a secondary process. With regard to the distribution of class and person, the EC languages differ considerably (cf. Magometov 1965:196f.): Lezgian and Agul have neither class nor personal inflection (Topuria 1959:10; Magometov 1970:44 refers to "Okamenelye Klassnye pokazateli v agul'skom jazyke"); in Tabasaran reduced class inflection is connected with more recent differentiation of persons (cf. Schmidt 1972); the inflection of the Udi verb is restricted to person (Žeiranišvili 1971:79ff.; Pančvize 1974:131 ff.; Schulze 1982:143ff.).

§3. Among the WC languages Abkhaz-Abaza (A.-A.) differentiates between *masculine rational* vs. *feminine rational* (2nd sg., 3rd sg. pronoun, 3rd sg. ergative) and *rational* vs. *non-rational* (3rd sg.) both in the absolute pronominal inflection and in the pronominal elements incorporated in the verb (cf. Deeters 1931:293; 1955:27; 1963:52f.; Lomtadze 1967:111, 113, 133, 135; Žirov/Ėkba 1956:605; Gecadze 1979:11ff.; Klimov/Alekseev 1980:30f.).

(9)

Person

	Class	Pers. Pron.		Pers. Prefix			
		Singular	Plural	Singular		Plural	
				Ergative	Absolute	Ergative	Absolute
1		sara	hara		s		h(a)
2	Masc. rat	wara	š°ara		w		š°
	Fem. rat.	bara			b		
3	Masc. rat.	jara	dara	j	d		
	Fem. rat.	lara		l		r	j
	Non-rat.	wara		a	j		

Although Rogava's (1953) attempt at identifying the personal prefixes of A.-A. with the EC CM cannot be regarded as successful (cf. Deeters 1955:28ff.), the A.-A. examples in (9) could well reflect indirectly older WC class inflection. This hypothesis is based on three arguments: (a) the parallel of an allomorph split of the ergative in a few EC languages (cf. (4)): in both cases the semantic differentiation of older class inflection (cf., in particular, I, II, III in (6)) was followed by the development of new, independent paradigms. As to A.-A., the process corresponds to Benveniste's definition of *innovating mutations*: "Innovating mutations result from the loss or emergence of formal classes, processes which thus modify the total stock of available categories (Benveniste 1968)." Therefore the A.-A. categories in (9) could well represent a WC archaism diachronically conditioned by older class inflection; (b) the fact that class inflection can be completely lost (cf. Lezgian, Agul, Udi); (c) the probable relationship between EC and WC, a theory that is based on word correspondences (cf. Trubetzkoy 1930). In this context the Proto-North Caucasian model preceding Proto-EC and Proto-WC would imply a common grammatical system which in the light of arguments (a) and (b) may be considered as class inflecting. From the semantic point of view, this Proto-North Caucasian reconstruction would confirm a series of archaisms which are still relevant for common EC itself: lack of number, early development of the categories human vs. non-human and masculine-rational vs. feminine-rational. But this hypothesis cannot be

proven without the formal identification of the Proto-North Caucasian CM.

§4. In contrast to WC, there is no evidence whatsoever for older class inflection in SC. The arguments in favor of noun classes put forward by Žavaxišvili 1937, Čikobava 1942; 1979 and Vogt 1942:255 have been rejected by Deeters 1955 and recently by Oniani 1985, who, however, does not discuss Žavaxišvili 1937. Žavaxišvili failed in his attempt to identify the anlaut of Georgian nouns referring to persons and animals with the EC class prefixes *w* (I) or *b* (III) on two counts: the interpretation of single items and the unsolved question of function. Čikobava distinguishes two classes, *man* and *thing*: “*a d a m i a n i s* ო ადამიანის კატეგორიას შეეძლება ჰქონდეს ‘vin?’, *n i v t i s a s* ო ნივთისას ჰქონდეს ‘ra?’, — გაირჩევა ‘vin,’-*კ ა ტ ე გ ო რ ი ა* და ‘ra’-*კ ა ტ ე გ ო რ ი ა*” (Čikobava 1979:89). This theory is not, however, confirmed by the evidence. As we have learned above all from the pioneering works of the great scholar we are honoring in these pages, what Čikobava identifies as prefixes for *man* (*m-*) and *thing* (*s-*, *d-*, *n-*) are elements of word formation, not of class inflection. Šanize 1973:135 cites formations in *me...e*, *me...uri* designating profession:

- (10) Georgian *me-urm-e* ‘driver, wagoner’ : *urem-i* ‘carriage, wagon, car’; *me-bağ-e* ‘gardener’ : *bağ-i* ‘garden’; *me-bad-uri* ‘fisher’ : *bad-e* ‘net’, etc.

formations in *sa...e*, *sa...o*, *sa...uri* designating purpose or determination:

- (11) Georgian *sa-katm-e* ‘hen-coop’ : *katam-i* ‘hen’; *sa-kveqñ-o* ‘public’ : *kveqñana* ‘world, country’; *sa-ta-uri* ‘title, heading’ : *tav-i* ‘head’, etc.

In the SC languages we observe verb doublets used in agreement with the *animate* or *inanimate* quality of the connected substantives (Klimov/Alekseev 1980:127ff.):

- (12) Georgian, animate : inanimate: *çola* : *deba* ‘to lie’; *çakceva* : *davardna* ‘to fall’; *daxrçoba* : *çazirwa* ‘to be drowned’; *moxuceba* : *dazveleba* ‘to grow old’, etc.

In contrast to (9), however, these examples cannot be interpreted as a reflex of the *innovating* mutations of a lost class inflection in the strict sense of the word. This conclusion would be disproved among other things by the semantics of the EC classes I, II, and III (cf. (6)) which stress the differences between *human* vs. *non-human* and masculine-rational (I) vs. feminine-rational (II). Klimov (Klimov/

Alekseev 1980:127) considers this archaic material as evidence for his hypothesis classifying SC as an active language.² It may well be, however, that the extremely old dichotomy *animate* vs. *inanimate* was relevant in the early stages of EC and WC as well (see (8)), although Klimov himself regards their reconstructions not as *active*, but as *ergative languages*.

Universität Bonn

Notes

⁰I would like to thank Karin Hlaváček and Dr. Walter Fitzgerald for correcting my English, and Dr. Wolfgang Schulze for the correction of a Chechen translation.

¹Magometov 1970:44 refers to "okamenelye klassnye pokazateli v agul'skom jazyke."

²Although I regard Klimov's ideas (cf. Klimov/Alekseev 1980:127ff.) as very stimulating, I disagree with his explanation of SC as an actual *nominative-accusative* language (!) going back to an older *active* language.

References

- Benveniste, E. 1968. "Mutations of Linguistic Categories." *Directions for Historical Linguistics. A Symposium*, ed. by W. P. Lehmann and Y. Malkiel, 83-94. Austin, Texas.
- Bokarev, E. A. and K. V. Lomtadze. 1967. *Jazyki narodov SSSR, tom IV: Iberijsko-kavkazskie jazyki*. Moskva.
- Cercvadze, I. S. 1967. "Andijskij jazyk." In Bokarev/Lomtadze (edd.), 276-92.
- Cercvaze, I. 1965. *Andiuri ena*. Tbilisi.
- Čikobava, A. 1942. *Saxelis puzis uzvelesi agebuleba kartvelur enebši*. Tbilisi.
- _____. 1979. *Iberiul-kaŋkasiuri enatmecnierebis šesavali*. Tbilisi.
- Čikobava, A. and I. Cercvaze. 1962. *Xunzuri ena*. Tbilisi.
- Deeters, G. 1931. "Der abchasische Sprachbau." *Nachrichten von der Ges. d. Wiss. zu Göttingen. Phil.-Histor. Klasse*, 3,2:289-303. Berlin.
- _____. 1955. "Gab es Nominalklassen in allen kaukasischen Sprachen?" *Corolla Linguistica. Festschrift für F. Sommer*, 26-33. Wiesbaden.
- _____. 1957. "Die Stellung der Kharthwelsprachen unter den Kaukasischen Sprachen." *Revue de Kartvélogie Bedi Kartlisa* 23:12-16.
- _____. 1963. "Die kaukasischen Sprachen." *Handbuch der Orientalistik*, ed. by B. Spuler, I. Abt., 7:1-79. Leiden-Köln.
- Žavaxišvili, I. 1937. *Kartuli da kaŋkasiuri enebis tavdaŋirveli buneba da natesaoba*. Tbilisi.
- Žeiranišvili, E. 1971. *Udiuri ena*. Tbilisi.
- Gecadze, I. O. 1979. *Očerki po sintaksisu abxazskogo jazyka*. Leningrad.

- Gudava, T. E. 1967. "Bitlixskij jazyk." In Bokarev/Lomtatidze (edd.), 293-306.
- Hockett, C. F. 1964. *A Course in Modern Linguistics*. 7th ed. New York.
- Jakovlev, N. F. 1940. *Sintaksis čečenskogo literaturnogo jazyka*. Moskva-Leningrad.
- Ķaxaze, O. 1979. *Arčibuli ena da misi adgili dagestnis monatesave enata šoris*. Tbilisi.
- _____. 1984. *Gramatīkuli ħlasebi lezgiur enebši*. Tbilisi.
- Kibrik, A. E., S. V. Kodzakov, I. P. Olovjannikova and D. S. Samedov. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka I, II, III*. Moskva.
- Klimov, G. A. and M. E. Alekseev. 1980. *Tipologija kavkazskix jazykov*. Moskva.
- Kuryłowicz, J. 1964. *The Inflectional Categories of Indo-European*. Heidelberg.
- Lomtatidze, K. V. 1967. "Abxazskij jazyk." In Bokarev/Lomtatidze (edd.), 101-122.
- _____. 1967. "Abazinskij jazyk." In Bokarev/Lomtatidze (edd.), 123-144.
- Magometov, A. A. 1965. *Tabasaranskij jazyk*. Tbilisi.
- _____. 1970. *Agul'skij jazyk (Issledovanie i teksty)*. Tbilisi.
- _____. 1982. *Megebskij dialekt darginskogo jazyka*. Tbilisi.
- Oniani, A. L. 1985. "O grammatičeskoj kategorii klassa v kartvel'skix jazykax." *VJ* 1985/3:67-76.
- Pančvize, V. 1974. *Uduri enis gramatīkuli analizi*. Tbilisi.
- Rogava, G. V. 1953. "K voprosu o perexode klassnogo sprjaženija v ličnoe sprjaženie v iberijsko-kavkazskix jazykax." *Soobščeniya AN GSSR* 14:443ff. [As this article was not available to me, I cite it after Deeters 1955:28⁶.]
- Šanize, A. 1973. *Kartuli enis gramatīkis sapuzvlebi. I. Morpologia*. 2nd ed. Tbilisi.
- Schmidt, K. H. 1972. "Probleme der Typologie (Indogermanisch/Kaukasisch)." *Homenaje a Antonio Tovar*, 449-54. Madrid.
- Schulze, W. 1982. *Die Sprache der Uden in Nord-Azerbajdžan. Studien zur Synchronie und Diachronie einer süd-ostkaukasischen Sprache*. Wiesbaden.
- Timaev, A. D. 1983. *Kategorija grammatičeskix klassov v naxskix jazykax*. Rostov.
- Topuria, G. 1959. *Lezgiuri zmnis ziritadi morpologiuri ħategoriebi*. Tbilisi.
- Trubetzkoy, N. S. 1930. "Nordkaukasische Wortgleichungen." *WZKM* 37:76-92.
- Vogt, H. 1942. "Varia. La parenté des langues caucasiqes." *NTS* 12:242-57.
- Wackernagel, J. 1926; 1928. *Vorlesungen über Syntax*. I; II. Basel.
- Xajdakov, S. M. 1967. "Arčinskij jazyk." In Bokarev/Lomtatidze (edd.), 608-26.
- _____. 1980. *Principy imennoj klassifikacii v dagestanskix jazykax*. Moskva.
- Žirov, X. D. and N. B. Ėkba (edd.). 1956. *Russko-abazinskij slovar'*. Moskva.

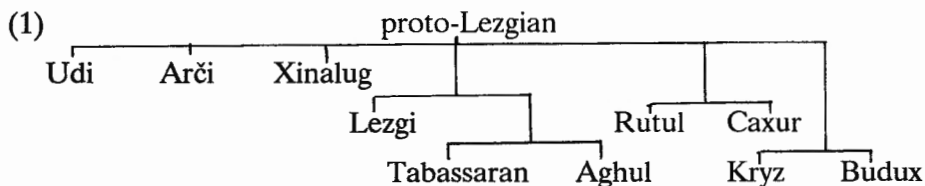
Tracing Aspect Coding Techniques in the Lezgian Languages

Wolfgang Schulze-Fürhoff

The aim of this paper is to discuss some features of aspect coding techniques in the south-east Caucasian languages. It is part of a first approach towards a comparative grammar of these languages, also called "Lezgian languages." I will confine myself to some morphosyntactic problems, as the lack of texts, especially those regarding the smaller languages, does not allow a more content-oriented approach at the present. Furthermore, it is precisely the technique of encoding aspectual differentiations which may give us hints for both a typological and a diachronic classification of the languages in question. This classification can lead to some assumptions on possible coding techniques in the proto-Lezgian language.

1. Introduction. The ten Lezgian languages constitute the southern branch of the East Caucasian languages (ECL). Though objections have recently been raised to the inclusion of Xinalug in this group (cf. Alekseev 1984), I shall keep to my classification of this group given in Schulze 1984 until we have stronger evidence that a reinterpretation is necessary.

Thus the Lezgian languages comprise ten single languages, spoken both in close and scattered areas in southern Daghestan.¹ A tentative genetic classification draws the following picture:



I do not want to discuss this classification in more detail here. It may be used as a guideline for the interpretation of those morphosyntactic phenomena that will be introduced in this paper.²

As Comrie has stated (1976:14): "there are relatively few works on aspect from the viewpoint of general linguistics, rather than in individual languages." Though Comrie's study itself has obviously improved this

situation, we have to stress that there is still a lack of theoretically oriented work on this matter, especially regarding the results of modern theories of syntax.

It is not the objective of this paper to fill this gap. Thus I do not want to enter into the old controversy on the question of what role is played by the so-called aspect opposition in a given system, or even whether the assumption of such an opposition is true or not.³ I will confine myself to some general remarks that are necessary in order to evaluate the data of the Lezgian languages.

As for a definition, I refer to the quite general remarks by A. Meillet: "La catégorie de l'aspect (...) embrasse tout ce qui est relatif à la durée et au degré d'achèvement des procès indiqués par les verbes" (Meillet 1921:183). It is clear that this description of what "aspect" is still shows some confusion between "aspect" and "Aktionsart," but it may work as an introduction into what is called "aspect." A perhaps more precise description is given by H. Bußmann in her *Lexikon der Sprachwissenschaft*: "Grammatische Kategorie des Verbs, durch die die subjektive Perspektive ausgedrückt wird, unter der der Sprecher den in der Verbalform bezeichneten Vorgang in Relation zur Sprechsituation sieht" (Bußmann 1983:46). Thus the phenomenon called "aspect" is directly connected with the speaker's relation to a reported act, whereas "Aktionsart" refers to the grammatical methods by which the reference to special attributes of the action is encoded. In other terms, "aspect" has something to do with connotation, whereas "Aktionsart" implies denotation. Such an observation had already been made by E. Hermann in 1927. In his article "Objektive und subjektive Aktionsart" (Hermann 1927) he differentiates between an objective description of an action, which can be durative or nondurative, and a subjective one, both "complexive" or "cursive." The "cursive" and "durative" interpretation both share the feature "imperfective," whereas "complexive" and "nondurative" share the feature "perfective." In other terms, Hermann refers to the difference "point" and "ligne," which can be formalized as:

- (2) A _____ B (ligne) A.B (point)
(cf. Hermann 1927:212).

This basic differentiation is often elaborated in individual languages. Thus we have numerous ways of representing aspectual content, ways that are reflected by quite different terminologies. In his 1962 reader Maslov quotes such terms as "finite/non-finite," "telic/atelic," "perfective/non-perfective," "definite/non-definite" (Maslov 1962:14). It is important to see that these terms always reflect a binary opposition,

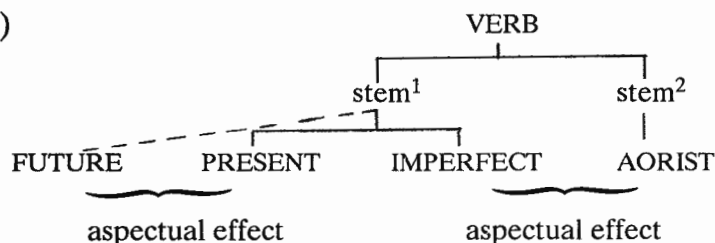
though this is not true for all languages, cf. the three aspects in Greek "duratif," "achevé," and "perfectif," as they are described by Meillet.

Comrie 1976:25 has demonstrated that the "imperfective" aspect involves secondary aspectual categories such as "habitual" vs. "continuous," and, depending on the latter, "progressive" vs. "non-progressive." Regarding the situation in the Lezgian languages, such fine work can hardly be done. On the one hand most of the grammarians either neglect this problem or use such vague terms as the Russian "sovershennyj/nesovershennyj" or "sruli/usruli" in Georgian, on the other hand the textual data for most languages are much too small to allow a precise description of each connotation. Without further investigation we have to settle for the use of such terms as "perfective/imperfective" or "resultative/non-resultative."

Before turning to the data of the Lezgian languages it is necessary to mention another quite important feature of aspect differentiation. When we claim that "aspect" refers to the subjective interpretation of a given action, it becomes clear that it is directly connected with witnessed and nonwitnessed actions. A witnessed action allows the speaker to concentrate on how the action has been performed, thus emphasizing the "line of action," whereas a nonwitnessed action reported by the speaker tentatively stresses its end result, i.e., its perfectiveness. This inferential quality of the perfective aspect, that is, its orientation to the present interpretation of the speaker leads to the fact that it is often expressed (or even substituted, cf. Meillet 1921) by the perfect tense, cf. the definition of the perfect tense given by J. P. Vet in a review of Comrie 1976: "(...) the perfect merely expressed the fact that one considers a situation from a reference point that is later in time" (Vet 1977:384).

This point leads us to the relation of the two categories "time" and "aspect." It has been claimed for the history of many languages that their protolanguage originally had only an aspect differentiation, which merged with a tense differentiation later on. This may be true for the Indo-European protolanguage, but is difficult to prove for the Lezgian languages. Nevertheless, nearly every Lezgian language shows something that looks like an aspect differentiation, but the categories often fade within the system of time expressions.⁴ In some instances however there is a tendency to use different verbal stems as the basis for the individual tenses, which reflect aspectual features. This phenomenon can be formalized as follows:

(3)



One of the major questions concerning aspect coding techniques in the Lezgian languages is whether there is a common way of differentiating the two (or more) basic stems, in order to draw a common (perhaps proto-Lezgian) denominator from these data.

2. Types of aspect differentiation. J. Holt 1943:45 has introduced two major types of aspect coding techniques. He speaks of "aspect flexionelle" and "aspect derivative," terms that should be replaced by those proposed by Comrie 1976, namely by "morphological aspect" and "syntactic aspect." The first is characterized by its synthetic, the latter by its analytic makeup. Both types can be found for the Lezgian languages, but I want to concentrate on the "morphological aspect" only. As to coding techniques, this type can be subdivided into various subtypes. I shall adhere to the following classification: a morphologically marked aspect can be encoded by means of 1. prefixation, 2. internal change, 3. suffixation, 4. suppletion. Prefixation, a method that is well-known from Georgian, Russian, etc., is quite rare in the Lezgian languages. It appears systematically only in the southern dialect of Tabassaran (ST). Quite analogically to the above mentioned languages, a perfective stem is formed from the imperfective one by means of prefixes, which originally denoted specific types of direction, cf.:⁵

- (4) ST *ġuz* (imperf.) 'to carry,' *ġa-ġnu* (perf.)
 uġuz (imperf.) 'to go,' *ġ-uġnu* (perf.)

The prefix *ġV-* is only used with finite verbs; together with participles and gerunds a prefix *dV-* is used instead, cf. the two sentences:

- (5) ST *dumu ġ-uġ-nu ha-iz*
 he-ABS PERF-go-PRET wood-DAT
 'He went to the woods.'
- (6) ST *d-uġ-nu žaġ ġ-ap'-nu*
 PERF-go-PRET call-ABS PERF-make-PRET
 'gone he made a cry'
 'Having left, he made a cry.'

Though the northern dialect of Tabassaran also tends to use such prefixes — but only to emphasize the perfective aspect, which is also marked by other means — the data of southern Tabassaran are unique within the Lezgian group. This is astonishing, as nearly every language exhibits a more or less large corpus of local preverbs, which could have been used for aspect differentiation just as in Tabassaran. The question as to why the other languages do not behave like Tabassaran has not yet been investigated. We can assume that the reason may be either the tendency to neglect aspectual differences, as in Lezgian, or the existence of quite varying methods of aspect coding techniques, which have hindered a systematic use of such preverbs. Probably the most common technique of aspect differentiation in the Lezgian languages is that of internal change in the verbal root. Here again we have to distinguish two types: “ablaut” and “infixation.” The “ablaut” phenomenon, which has already been observed by Dumézil 1933, is well-known from Darginian, cf. Magometov 1982:72:

- (7) Darg. *baq’as* (perf.)/ *biq’es* (imperf.) ‘make’
 bak’as (perf.)/ *bik’es* (imperf.) ‘come’

It is interesting that within the Lezgian group only northern Tabassaran, which has close contact to Darginian (and Lak), has preserved this probably very old technique (cf. Schmidt 1968:190). Here again we have mostly to deal with an ablaut *a* (perfective)/ *i* (imperfective), cf.:

- (8) NTab. *žawk’-žiwk’* ‘wash’; *qārk-qīrk-* ‘return’

Another ablaut-type is that of *u/i*, which according to Magometov 1965:191 is very seldom found:

- (9) NTab. *uč^w-ič^w-* ‘go’; *uwš^w-iwš^w-* ‘mow’

Both types of ablaut share the common feature of vowel change from back to front. Such a change can be traced in some verbal roots in Lezgian too, though it does not exhibit a specific function in the present-day language (cf. Moor 1985:19). I do not want to give a full account of the very complicated methods that are used in Lezgian to distinguish the different verbal stems,⁶ but let me quote the following examples which show this old distinction quite clearly:

- (10) Lezg. *ak’un/ ek’iz* (absolute) ‘sink’
 žugun/ žigiz (absolute) ‘find’

Lezgian shows a tendency to generalize the stem vowel, using ablaut forms for the thematic vowel instead (cf. Moor 1985). Though both are now unproductive for aspect differentiation, it cannot be clearly

determined which vowel was the original basis for such a technique in an earlier stage of the language (cf. Moor 1985:20).

A change of the thematic vowel can also be observed in Aghul, a language, which in spite of its strong relationship to Tabassaran, does not show any trace of stem-vowel ablaut. In Aghul the thematic vowel changes quite systematically when the verbal root is used as a past gerund, cf. Magometov 1968:244; 1970:128:

- (11) Aghul. aq'-a-d (ger.pres.)/aq'u-n (ger.past) 'doing'
 xur-a-d (ger.pres.)/xur-u-n (ger.past) 'reading'

None of the other languages shows any systematic trace of such an old ablaut system. Nevertheless, there are two verbs in Udi which may reflect it in a lexicalized form, cf. Schulze 1982:168:

- (12) Udi aq'-sun / b-iq'-sun 'take'

The initial *b-* in *b iq'-sun* probably is a petrified class marker, the motivation of which is unclear (perhaps it results from the 'imperfective' action, drawing the listener's attention to the object more than to the result).

The second type of internal change is that of infixation. Once again it is Northern Tabassaran which gives the clearest examples. Here verbs of the so-called "second conjugation" insert a morpheme *-l-* just after the stem vowel in order to encode the imperfective aspect. There are numerous examples of such verbs, most of them having *-i-* as stem vowel. Thus we might think of a nearly complementary distribution of ablaut and infixation (the only counterexample I have found is *ulup-us/ululpun-di* 'show'). Let me quote as an example (Magometov 1965:193):

- (13) NTab. bik'-/bi-l-k'- 'write'
 birx-/bi-l-x- 'sow'
 qirx/qi-l-x- 'ask'
 t'iwx-/t'i-l-x- 'fly'

The last three examples show that the aspect marker tends to replace the class marker that appears with the perfective aspect (cf. Magometov 1965:192).

There are only few traces of such coding techniques in the other Lezgian languages. Surprisingly enough, it is precisely the most "demolished" Lezgian language, i.e., Budux, that gives us the clearest evidence for an infixation technique in the protolanguage. Though the language no longer has a clear aspect system, one of the methods to distinguish between the present stem (the basis for the masdar) is to reduce the number of phonemes in the root (cf. Mejlanova 1984:195):

- (14) Budux q'urot'u ____ q'ut' 'stress'
 sarqar ____ saq 'die'
 halsal ____ hasil 'want'
 q'arzar ____ q'azur 'rise'

With some exceptions (e.g., *sonk'an* / *sork'un* 'jerk'), either *-l-* or *-r-* is involved in this process. Both phonemes also appear as markers of the present stem, as the examples in (14) show. It is interesting to see that in some verbs these present stem elements are infixes just as in Tabassaran, and are drawn from the stem in the second stem, cf.:

- (15) Budux e-l-qi ____ eq-il 'sit'
 ha-r-ki ____ hak-ir 'soil'

We can assume that this procedure took place in two steps. First both stems were differentiated by the presence or absence of the morpheme *-l/-r-*, then it was reestablished under the pressure of verbs like those given in (14).

It is probable that such phenomena also exist in Kryz, which is directly related to Budux, but unfortunately we don't have any good grammatical account of this problem for Kryz. Still, the remark of Saadiev 1967:636, that Kryz too uses two different stems as the basis for the present and past gerunds may hint at a parallel differentiation.

Let us once again turn to Aghul, a language that does not exhibit the category of aspect (cf. Magometov 1970:156). The dialect of Burxikan still shows traces of infixation parallel to those in Budux. Some verbs mark the difference between present and past gerunds not only by means of the ablaut of the thematic vowel, as shown in (11), but also by means of the loss of an infix *-r-* in the past gerund, cf.:

- (16) Aghul/Burxikan a-r-q'a-ri ____ aq'u-na 'making'
 di-r-ža-ri ____ diž-u-na 'compiling'

These two different stems are also relevant for the formation of the imperative mode. Whereas the positive imperative is based on the past gerund, the prohibitive uses the infixes present gerund, a type well-known, for example, from Georgian and Classical Armenian,⁷ cf.:

- (17) Aghul/Burxikan aq'e 'make!' m-a-r-q'a 'don't make!'
 NTab. ap'in 'make!' m-i-p'an 'don't make!'
 Geo. c'ere 'write!' nu c'er 'don't write!'
 Arm. ber 'carry!' mi berer 'don't carry!'

Arči plays a considerable role in the investigation of internal aspectual change. Dumézil 1933:78 points out: "Dans un très grand nombre de

verbes, l'arči insère -r-, au présent seulement, toujours immédiatement devant la consonne qui est sentie.(...)" Dumézil hesitates to follow Dirr 1928:261, who regards those forms as probably durative, but Kibrik's investigations (Kibrik 1977) strengthen Dirr's view.

Indeed, Arči exhibits a distinctive aspectual opposition, which is based on both infixation and change of the thematic vowel. Moreover, most durative stems show the sonantic element -r as a suffix too, thus achieving a status quite similar to that of Budux. Still, the conditions under which both infixation and suffixation appear are rather complicated, though always complementary. In short we have to deal with features of stem structure, existence and place of stress, and reduplication, cf.:

(18) Arči	á-k'a _____	á-r-k'u-r	'fall'
	á-ti _____	á-r-ti-r	'search'
but	á'k: ^w a _____	á-k: ^w u-r	'see'
	ír: ^w in _____	ír: ^w in	'work'
	dárh <u>u</u> _____	dárh <u>u</u> -r	'think'

As for the majority of Arči verbs, they follow at least one of the restrictions for infixation given above. We can observe that the suffix -r plays the major role in encoding the durative aspect (cf. K'axaze 1979:501), but that there is also a strong tendency towards the lexicalization of infixed forms (cf. *ír:^win* 'to work'). Thus Arči follows Budux with respect to the fate of infixation.

Both Xinalug and Udi no longer show any trace of infixation. But there is a morpheme in each of the two languages which might be linked to the above-cited durative infix -r/l-, but, however, whose function is different.

The analysis is based on the repetition of the infix as a suffix together with a partial loss of the infix, as observed in Aghul and Budux. This process can be formalized as:

durative: VE-r/l-RB → VE-r/l-RB-r/l → VERB-r/l

Assuming that this process is a quite general tendency in languages which strengthen their agglutinating character (perhaps through language contact), it seems possible to connect the morphemes for the future tense in Xinalug and Udi with the old infix -r/l-.

Before turning to the individual data, it is necessary to take a brief look at the interaction of aspect and tense (mainly future tense). As we have seen, infixation always provokes the imperfectivization of a given perfective root (stem). This is also true for suffixation, whereas prefixation leads to the perfectivization of a given imperfective stem. We can draw the following scheme regarding the feature of markedness:

unmarked	marked	morphological means
PERFECTIVE	IMPERFECTIVE	INFIX/SUFFIX
IMPERFECTIVE	PERFECTIVE	PREFIX

Not only what has been said for Tabassaran, but also the data for the Kartvelian and Indo-European languages, demonstrate that the first opposition always precedes the second one diachronically. This claim is supported by the fact that we can most often trace the original (directive) function of prefixes, whereas the suffixes and infixes hardly allow any semantic assignment.

The old distinction between an unmarked perfective and a marked imperfective tended to be copied by a parallel tense opposition. The result is what is called 'aorist-stem' vs. 'present-stem' in Indo-European languages, or first and second series in Kartvelian languages. The present stem already shows the generalization of a secondary function that the so-called 'injunctive' in Indo-European possessed, namely the 'hic et nunc'-tense. Due to the developing tense differentiation the system required morphological means to express the imperfective aspect of the past (aorist) tense. This gap was either filled by the (augmented) injunctive or derived aorist forms (as in Slavic), or by adding secondary suffixes to the present stem (as in Kartvelian). The result is a well organized system of aspect and tense opposition:

aspect	PERFECTIVE	IMPERFECTIVE
tense	AORIST $\xrightarrow{\text{in/suffix}}$	"INJUNCTIVE" $\xrightarrow{\text{hic et nunc}}$ PRESENT
	AORIST $\xrightarrow{\text{in/suffix}}$	"INJUNCTIVE" $\xrightarrow{\text{or changed}}$ IMPERFECT

It is quite obvious that this system can easily be equated with the scheme given in (3). In order to reflect the interaction of present tense and imperfective aspect on the one hand, and aorist and perfective aspect on the other, it seems appropriate to follow a proposal of Šanize 1976:69 and Aronson 1982:41 in labelling it by means of the Georgian term *mc'k'rivi*, "screeve."⁹ Thus the two screeves exhibit the following features:

	screeve ¹	screeve ²
aspect	imperfective	perfective
tense	present imperfect	aorist

Such a stable system necessarily tends towards entropy. This is mainly due to the fact that the tense function supersedes the old aspect opposition, and that there is a need to express a future tense by other means, not just through the secondary function of a conjunctive mood (cf. Old Georgian), which might be ambiguous. Thus, the tendency is either to give up aspectual differentiation in favor of a complex tense system or to rebuild the old opposition, which mainly fills the gap of the dichotomy in the present tense. Because of their specific semantics, local (and directive) preverbs are the “natural” source of a practical morphological means of encoding such a “perfective present,” cf. Lithuanian *abù ėjo pàs mēšką* ‘both went towards the bear’ (imperf.) and *priē tō ėžero priėjo* ‘they went to the lake (and reached it)’ (Leskien 1919:188). This tendency to use preverbs to encode a secondary perfective aspect has been stabilized, for example, in Old Church Slavonic. Here the prefixed form fills the gap in the present tense, developing towards a mere future tense, cf. Leskien 1969:173: “Da der Redende bei einer perfektivischen Handlung die Vollendung, nicht aber den Verlauf im Sinne hat, der Moment der Vollendung aber, wenn von einer solchen Handlung die Rede ist, in der Zukunft liegt, kann diese Präsensform nicht die in der Gegenwart andauernde Handlung ausdrücken, sondern bekommt (...) den Sinn der zukünftigen Zeit.” An example of this usage is OCS: *ašte kǫto ljubitǫ mę, slovo moje sǫbljudetǫ i otčę moi vǫzljubitǫ i* ‘si quis diligit me, sermonem meum servabit et pater meus diliget eum’ (John 14:23).¹⁰ While the Kartvelian languages show a parallel tendency to use preverbs as perfective markers (cf. Svan. *asq’ę:ne* ‘he does’ vs. *an-sq’ę:ne* ‘he will do’ (Schmidt 1984:32), only Southern Tabassaran behaves similarly within the Lezgian group. The other languages have not been affected by this process up to now. But if we try to trace any old aspect markers within the morphological inventory of the future tense, we might expect to meet preverbal forms. As for the two languages in question, Xinalug and Udi, this is not true: the future tense in Xinalug is directly connected with the “non-resultative” stem (see below), to which the indicative morpheme *-mā* (or an augmented form *-zmā*) is added. The distribution of both morphemes is quite remarkable. *-mā* is used with imperfective stems in *-r(i)* only and “označæet dejstvie, kotoroe soveršitsja v buduščem” (Kibrik et al. 1972:173), cf. *yā oǰur dā pʰil hana lat:iq’-ir-mā* ‘then I will return the money to you.’ *-zmā* is used with all forms of imperfective stems, but together with *ri*-imperfectives it “oboznačæet dejstvie, kotoroe budet dlit’sja v buduščem bez ukazanija na zaveršennost’,” whereas it leads to perfectivization of all other stems, cf. *hini futbol anskʰi-r-zmā* ‘he will be playing football,’ but *čʰek’wisoǰ tʰotʰuǰ zi c’wa kwi-d-mā* ‘after the sale I will go home.’ Thus

we can observe that in the future tense, the suffix *-mä* behaves like preverbs in other systems, i.e., it leads to the perfectivization of an imperfective (present) stem (cf. *k^hirağ kiza gāši-ri* 'today it is snowing'). In order to mark a general, nonconcrete future, the morpheme *-z-* is added. It is obvious that the original morpheme for the imperfective was the suffix *-ri*, which later entered a new type of aspectual paradigm (cf. below). This morpheme can easily be connected with the *l/r*-imperfective of the other Lezgian languages.

Udi, just as Xinalug, has two different future tenses. However, they do not reflect an aspectual opposition, cf. the Udi version of John 14:23 (see above): *wā šute but'uq'sa zaṣ, šet'in tamnebesa* (PRES) *bez aitaṣ wā bez babal* (sic!) *buq'al-t'u* (FUT¹) *šot'u*, and Mk 3:23 *etärt'ubako* (FUT²) *šeitana č'ewk'es šeitanaṣ* 'Pôš dýnatai satanâs ekbállein.' The two morphemes *-al* (FUT¹) and *-o* (FUT²) tend to be interchangeable, though specific data show that the *al*-future denotes a general, nonconcrete tense which can merge with the present tense, whereas *-o* encodes a futurum exactum (Schulze 1982:157). The fact that *-al* is also used as the morpheme for the participle of simultaneousness strengthens the argument for interpreting it as the old Lezgian imperfective morpheme *-l/r-*, which took on the function of a future morpheme when first the old aspectual opposition was lost and then a new present tense (based on the masdar, cf. *arcesun* 'sit,' *arcesa* 'sitting') developed.

Bearing in mind, that infixation can also be traced in Rutul (cf. *ku-l-qw-ara* 'cut' - *ku-quri* (< **ku-q^w-uri*, Dumézil 1933:83, Dirr 1912:70)), we can speak of a formerly general technique for marking the durative aspect. The changes *r > l* and *l > r* are well attested in these languages, but it is hardly possible to determine which of the two morphemes was the original one.

Before taking a brief look at suppletive forms we have to mention another type of aspect differentiation, i.e., the use of different suffixes, partly restricted to single tenses. The clearest example is given by Xinalug (Kibrik et al. 1972:89). Here both aspects (called "resultative" and "non-resultative") are marked by different suffixes. There are at least four conjugation types in Xinalug, constituted by different sets of aspect markers, cf.:

(19)	NON-RESULTATIVE	RESULTATIVE
I	-iri	-wi
II	-ri	-i
III	-ndä	-ni
IV	-li/-zi/-ri	-∅

We do not find any hint of internal change in the root by means of ablaut or infixation. Typologically similar to the other languages is the restriction of the two aspects to different sets of tenses. Though the morphemes *-ri* and *-i* are of proto-Lezgian origin, we have to think of the Xinalug system as one which has been developed at a time when the language already had separated from the protolanguage. The following examples illustrate the Xinalug aspect pairs (cf. Kibrik et al. 1972:103-5):

(20) Xinalug	I	k ^h -irí/ k ^h -wí	‘do’
	II	yíbi-ri/yíb-i	‘extinguish’
	III	ǰǰki-ndǰ/ǰǰki-ni	‘laugh’
	IV	ki-zí/ki	‘burn’

Finally, Udi is the only language which has developed an aspectual opposition confined to one tense, namely the past tense. Here again we have two different suffixes, placed directly behind the verbal stem, cf. (Schulze 1982:166-67):

(21) Udi	ar-i	(imperf.)/ar-e	(perf.)	‘came’
	p-i	(imperf.)/p-e	(perf.)	‘said’
	bak-i	(imperf.)/bak-e	(perf.)	‘became’

This system surely is of quite recent origin. Udi no longer shows the typical Lezgian formation of present and past stems. But as regards the morphological inventory, the imperfective suffix *-i* must be regarded as belonging to the proto-Lezgian stock of tense/aspect marking morphemes. It is used in Udi to mark secondary imperfective tenses, even on the basis of the perfective, cf.:

(22) Udi	PRES	bak-s-a	‘become’
	IMPERF	bak-s-a-i	
	AOR	bak-i	
	PERF	bak-e	
	PRET	bak-e-i	
	FUT ²	bak-o	
	OPT	bak-o-i	

The tentative “imperfective” function of the suffix *-i* is reflected in several Lezgian languages, for instance in Lezgian, Rutul, and Caxur (Žeiranišvili 1984:575).

The other way of expressing an aspect opposition is by means of stem suppletion. The basic assumption is that some verbs have a very specific aspectual connotation which does not allow the speaker to use any systematic method in order to change its connotation without creating a paradox. The Lezgian languages exhibit such suppletive pairs in varying

degrees. In Udi this opposition sometimes reflects the earlier, now lost, categories “present” and “past gerund,” cf. (Schulze 1982:150):

- (23) Udi eġ- (PRES), ar- (PAST) ‘go’
 eġ- (PRES), p- (PAST) ‘say’

Xinalug, which still has these two categories, has two suppletive verbs (cf. Kibrik 1972:105):

- (24) Xinalug q’i (RES), kwi (NONRES) ‘be’
 x̣i (RES), kwi (NONRES) ‘go’

It is interesting to see that in both languages these two verbs are used as the basis for compound verbs. This may hint at a very archaic, semantically motivated differentiation. Contrary to Udi, the suppletive pairs of Xinalug are embedded in the aspectual system. The same can be supposed for an earlier stage of Lezgian, where the suppletive forms follow the differentiation of infinitive and absolute (cf. Moor 1985:1), cf.:

- (25) Lezgian t’ün (INF), nez (ABS) ‘eat’
 atun (INF), q:ʷez (ABS) ‘go’
 awun (INF), iġiz (ABS) ‘make’

Finally, this suppletion is very important in the case of the Lezgian auxiliaries. The data of the individual languages lead to two different verbs, cf.:

- (26) proto-Lezgian *wu- ‘be’ / *q’a- (IMPERF)

In applying Comrie’s terminology, we can say that the root *wu- represents the “absolute state” (cf. Spanish *ser*), and *q’a- denotes a “contingent state” (cf. Spanish *estar*). As both auxiliaries are used to express compound tenses in the Lezgian languages, it is probable that they produce something like a syntactic aspect differentiation in those languages that share this feature.

3. Conclusions. In coming to a final conclusion on what the Lezgian material offers we have to bear in mind that with the exception of Northern Tabassaran, Arċi and Xinalug hardly any language shows a true aspect system. We have to deal mostly with a situation parallel, e.g., to Avar, as it has been characterized by Charachidzé 1981:138: “Il y a certes des phénomènes aspectuels, ce que l’on pourrait appeler des ‘effets d’aspect’, mais ils ne s’organisent pas en série et non font pas l’objet d’un choix obligatoire.”

We have deliberately passed over such “aspectual effects,” which are mainly expressed analytically, as they are of very recent origin in each

language and thus can hardly give any hint of a proto-Lezgian aspect system. Furthermore, the analytic forms are directly connected with the huge number of tenses that we encounter in some Lezgian languages. This question calls for a separate study.

In discussing the morphological coding techniques, we find that the prospects for retracing a proto-Lezgian system are much improved. The following table may serve as an overview of what can be found in the individual languages (productive or not):

(27)	prefixation	internal change	suffixation	suppletion
Lezgian	-	+	-	+
NTabas.	+	+	-	+
STabas.	+	-	-	+
Aghul	-	+	+	+
Rutul	-	+	+	+
Caxur	-	-	+	+
Arči	-	+	+	+
Kryz	-	?	+	+
Budux	-	+	-	+
Udi	-	?	+	+
Xinalug	-	-	+	+

The conclusions we can draw from this table are: 1. Suppletion is known in each language, though the verbs in question are not always the same. This type surely reflects a proto-Lezgian system. 2. Prefixing is a quite recent technique, known only in Tabassaran, systematically restricted to the southern dialect. 3. Suffixing is well documented, but the means are mostly confined to the individual languages. An exception is the general imperfective morpheme *-i*, which can be traced back to the proto-language. 4. Internal change probably is the oldest way of aspect differentiation. Its functionality in Northern Tabassaran which is preserved through contact with Darginian (and Lak), and its traces scattered all over the area of Lezgian languages demonstrate that this technique was the original in the proto-Lezgian language. It is not possible to say whether this was ablaut or infixation or even both.

University of Bonn

Notes

¹For a detailed discussion of what I have called "areal" and "village" languages see Schulze 1984.

²This classification does not reflect the specific positions of Arči, Udi, and Xinalug, as present-day data are much too uncertain to give a full account.

³Cf. Weinreich 1964:152 "Dann gibt es (...) sprachliche Aspekte überhaupt nicht, und wir haben die Aspektlehre ohne Rest aus der Sprachwissenschaft zu vertreiben."

⁴Cf. Žeiranišvili 1984:575.

⁵Cf. Schmidt 1968; Magometov 1965:190, 223.

⁶For a comprehensive study on the Lezgian verb see Moor 1985.

⁷Jensen 1959:94, 100; Tschenkéli 1958:186-87; Magometov 1965:276-77; Schmidt 1968:211.

⁸Arči, just as Xinalug, offers a system of two aspectual pairs, one for indicative, and one for conjunctive mood. In this paper I concentrate on the indicative mood only.

⁹Cf. Holisky 1981:3: "The term 'screeve' rather than the more familiar 'tense' is used to refer to a paradigmatic set of forms differing only in person and number (...)."

¹⁰Cf. the corresponding Lithuanian translation of John 14:23: *Jėi kàs mýli manė, jįs laikýs mào žodį, iř mào Tėvas jį mylės.*

References

- Anderson, J. 1973. *An Essay Concerning Aspect*. The Hague, Paris.
- Alekseev, M. E. 1984. "K voprosy o klassifikacii lezgingiskix jazykov." VJa 5 (1984):88-94.
- Aronson, H. 1982. *Georgian. A Reading Grammar*. Columbus, Ohio.
- Bußmann, H. 1983. *Lexikon der Sprachwissenschaft*. Stuttgart.
- Charachidzé, G. 1981. *Grammaire de la langue avar*. Paris.
- Comrie, B. 1976. *Aspect. An introduction to the study of verbal aspect and related problems*. Cambridge.
- Dešerieva, T. I. 1979. *Issledovanie vido-vremennoj sistemy v naxskich jazykach*. Moskva.
- Dirr, A. 1912. "Rutul'skij jazyk." SMOMPK 42,3:1-260.
- _____. 1928. *Einführung in das Studium der kaukasischen Sprachen*. Leipzig.
- Dumézil, G. 1933. *Recherches comparatives sur le verbe caucasien*. Paris.
- Giginejšvili, B. K. 1959. "O kategorii aspekta v udinskom jazyke." IKJ 9:361-84.
- Hermann, E. 1927. "Objektive und subjektive Aktionsart." IF 45:207-28.
- Holisky, D. A. 1981. *Aspect and Georgian Medial Verbs*. Delmar, New York.
- Holt, J. 1943. *Études d'aspect*. København.
- Ibragimov, G. X. 1978. *Rutul'skij jazyk*. Moskva.
- Jensen, H. 1959. *Altarmenische Grammatik*. Heidelberg.
- K'axaze, O. 1979. *Arčibuli ena da misi adgili dağest'nis monatesave enata šoris*. Tbilisi.
- Kibrik, A. E. et al. 1972. *Fragmenty xinalugsogo jazyka*. Moskva.
- _____. 1977. *Opyt strukturnogo opisaniya arčinskogo jazyka. T. II. Taksonomičeskaja grammatika. T. III. Dinamiceškaja grammatika*. Moskva.
- Leskien, A. 1919. *Litauisches Lesebuch*. Heidelberg.

- _____. 1969. *Handbuch der albulgarischen Sprache*. Neunte Auflage. Heidelberg.
- Magometov, A. A. 1965. *Tabasaranskij jazyk. Issledovanie i teksty*. Tbilisi.
- _____. 1968. "Kategorija vremeni v agul'skom jazyke." *IKJ* 16:231-45.
- _____. 1970. *Agul'skij jazyk. Issledovanie i teksty*. Tbilisi.
- _____. 1982. *Megebskij dialekt darginskogo jazyka*. Tbilisi.
- Maslov, Ju. 1962. *Voprosy glagol'nogo vida*. Moskva.
- Meillet, A. 1921. *Linguistique historique et linguistique générale*. Paris.
- Mejlanova, U. A. 1984. *Buduxsko-russkij slovar'*. Moskva.
- Moor, M. 1985. *The Lezgian verbal system*.
- Saadiev, Š. M. 1967. "Kryzskij jazyk." *Jazyki narodov SSSR*, 4:627-42.
- Šanize, A. 1976. *Zveli kartuli enis gramat'ik'a*. Tbilisi.
- Schmidt, K. H. 1968. "Aspekt und Tempus im Tabassaranischen." *Bedi Kartlisa* 25:208-13.
- _____. 1984. *Svan*. Manuscript (to appear in ILC I, New York).
- Schulze, W. 1982. *Die Sprache der Uden in Nordazerbajdžan. Studien zur Synchronie und Diachronie einer süd-ostkaukasischen Sprache*. Wiesbaden.
- _____. 1984. "Remarks on the Lezgian Case System." *Folia Slavica* 7:302-21.
- Talibov, B. B. and M. Gadžiev. 1966. *Lezginsko-russkij slovar'*. Maxačkala.
- Tschenkéli, K. 1958. *Einführung in die georgische Sprache. I*. Zürich.
- Vet, J. P. 1977. Review of B. Comrie "Aspect." In *Lingua* 41:382-85.
- Weinreich, H. 1964. *Tempus*. Stuttgart.
- Zeiranišvili, Evg. 1984. *C'axuri da mužaduri enebi. II. Morpologia*. Tbilisi.

**The Correspondence: Scythian Βαστακας
= Ossetian *bästä***

D. Testen

The entry for *bästä* in Abaev's *Istoriko-ëtimologičeskij slovar' osetinskogo jazyka* gives this word a translation 'country, world, region, place'. In his account of the history of the word, the author adduces a Scythian personal name Βαστακας, known through inscriptions from Greek settlements along the Black Sea coast: from Panticapaeum Βαστακας Μορδο[υ] and Δημοστρατε υιε Β[ασ]τακου, and from Tanais Μακαριον Βαστακου. This name is to be related to the Ossetian derived adjective *bästag*, which Abaev translates as 'local'.

The correspondence between Βαστακας and *bästag* was first presented by Vsevolod Miller. Vasmer (1925:35) expressed his doubts about its correctness rather succinctly ("...Kaum nach Miller *ŽMP*r 1886 Okt. 246 zu osset. *bästag* 'örtlich'."), but, as Zgusta notes, it is clear that Vasmer was troubled by the meaning 'local' as the basis for a personal name (Zgusta 1955:84). However, Zgusta observes that the more recent Miller-Freiman dictionary gives *bästag* a meaning 'dem Lande angehörend, örtlich, Landsmann', and concludes that Vasmer's reservations about accepting the correspondence are groundless. Abaev's first remarks on Βαστακας gloss *bästag* as both '*mestnyj*' and '*zemljak*' (Abaev 1949:185), although in his dictionary he states simply '*mestnyj*', evidently untroubled by any question of semantics which may have bothered Vasmer.

Whatever the meaning of *bästag* and its putative cognate Βαστακας may be, it is clear that the basic form *bästä* is to be taken to refer to a land or region. Abaev's etymological account for *bästä* concludes that the word goes back to an earlier **upasta-*, a form which has left no reflexes in the remaining Iranian languages, but which may be compared to Old Indic *upás-tha-* 'lap' (from *upás-* 'lap') which Abaev notes as having a secondary meaning 'abode'.

The sense of *upástha-* which Abaev mentions, 'abode', itself seems to be quite secondary; it is not to be found in any of the Vedic dictionaries. It must nevertheless be his claim that, if *upástha-* is to be taken to be a cognate of *bästä*, the development of 'lap' to 'abode' must be traced all

the way back to common Indo-Iranian, even though the Iranian sphere has entirely lost the etymon, with the putative exception of *bāstā*.¹

Is there not, however, a form nearer at hand than literary Sanskrit in which to look for a cognate to Ossetian *bāstā*? There is, in fact, a quite familiar Old Iranian term *upastā-*, which, in phonological terms, will yield an Ossetian reflex *bāstā* just as easily. Indeed, although the Ossetian *Auslautgesetze* are in many respects still unclear, a final *-ā* would seem preferable to short *-a(h)* as a source for final *-ä*. The problem is that *upastā-* has traditionally received a gloss ‘help, assistance’. The word has been analyzed as the particle *upa-* ‘below, under’ and the verbal stem *stā-* ‘stand’, and discussions of the meaning frequently give the structure of the word as comparable to German *Beistand*. The gloss has evidently been devised on the basis of the familiar phrase *upastām bara-*, found in both Old Persian and Late Avestan.

...pasāva adam Auramaz(d)ām patiyāvahyaiy Auramaz-
dāmaiṃ upastām abara...

Thereupon I sought aid from Ahura Mazda; Ahura Mazda bore me *upastā-*... (Darius Behistun I:54-5)

arəvdi sure anahite mošu mē java avanhe nurəm mē bara
upastām.

O Arvdi Sura Anahita, hasten to me at once with aid, bear me now *upastā-*! (Yašt 5:63)

It is significant that the verb employed in both cases is *bara-*. While translations of these passages have typically emphasized the aspect of ‘bringing’, the meaning of *bara-* is strictly speaking simply ‘bear, carry’. It may be instructive to compare the line from Behistun cited above with a similar line from the same text.

Auramazdā xšačam manā frābara

Ahura Mazda bestowed the kingdom upon me. (Darius Behistun I:12 passim.)²

Here, where the “giving” notion of the verb is evident, *bara-* has received the prefix *fra-*. One might suspect, therefore, that there is a semantic distinction between *bara-* and *fra-bara-* comparable to what we see in such pairs as Avestan *ay-/fra-ay-* (‘go/go forth’), *ar-/fra-ar-* (‘move/come forth’), *vaz-/fra-vaz-* (‘travel/go forth’), and *stā-/fra-stā-* (‘stand/step forth’), in which the sense of directionality is supplied by the preverb. If this

holds true for the pair *bara-/fra-bara-*, one might be tempted to see in simple *bara-* a verb meaning merely ‘carry, support’.

We may return now to the question of the noun *upastā-*. The situations in which the phrase *upastām bara-* is employed in the Behistun inscription of Darius I seem quite straightforward: Darius finds himself in a difficult situation — an usurping pretender, a crucial battle — Ahura Mazda bears him *upastā-*, and the problem is resolved in Darius’ favor.

We can get another perspective on Old Iranian *upastā-*, however, if we return to Ossetian. As has already been noted, Ossetian *bästā* works out quite well phonologically as a correspondent to *upastā-*. It is in the realm of semantics that rather formidable objections may be raised against lining up the two forms. There is, however, another Ossetian word *bāston* — by form, it would appear, an adjective derived from *bästā*. The meaning of *bāston*, nevertheless, is not ‘of a land or region’, but ‘basic, fundamental’. This suggests the prior existence of a form **bästā* with a sense of something on the order of ‘base’. That this was not an Ossetian innovation is evidenced by Middle Persian *apastāk* ‘foundation’, most familiar as a borrowing into Western scholarship in the form *Avesta*, referring to the fundamental text of Zoroastrianism.

The *bäst-* of *bāston* and the *apastā-* of *apastāk* would go back to an Old Iranian **upastā-*, a form identical to the one we have seen given the gloss ‘help’. If, however, we replace this with the sense extracted from the Ossetian and Middle Persian, and adjust the verb from ‘bring’ to non-directional ‘bear’, the Behistun line will become something on the order of “Ahura Mazda supported my foundation.”

If we return to the Behistun text and insert this alternative rendering for *Auramazdāmaiy upastām abara*, we will notice that this sentence is used in a rather specific set of situations: Ahura Mazda bears Darius *upastā-* on those occasions when the King’s right to rule is threatened or challenged. The phrase first occurs at the beginning of Darius’ reign, when Darius, with a few coconspirators, assassinates the usurper Gaumata, the Pseudo-Smerdis, and takes over the throne.

Saith Darius the King: Ahuramazda bestowed the kingdom upon me, Ahuramazda bore me [*upastā-*] until I got possession of this kingdom; by the favor of Ahuramazda I hold this kingdom.... (Darius Behistun I:24-6)

Saith Darius the King: There was not a man, neither a Persian nor a Mede nor anyone of our family, who might make that Gaumata the Magian deprived of the kingdom.... Not anyone dared say anything about Gaumata the Magian, until I came. After that I besought help of Ahuramazda;

Ahuramazda bore me [*upastā*]; of the month Bagayadi X days were past, then I with a few men slew that Gaumata the Magian, and those who were his foremost followers.... I took the kingdom from him. By the favor of Ahuramazda I became king; Ahuramazda bestowed the kingdom upon me. (Darius Behistun I:48-61)

The new king is then confronted by a series of rebellions, instigated by opportunists claiming (falsely, Darius carefully notes) to be the legitimate heirs of Cyrus, the founder of the Persian Empire, or of one of the dynasties overthrown by the Persians. Each of these Darius confronts successfully in a sequence of campaigns in which the phrase *upastām bara-* invariably figures prominently in the description of the battles.

Saith Darius the King: After that I went off to Babylon. When I had not arrived at Babylon, a town by name Zazana, beside the Euphrates — there this Nidintu-Bel who called himself Nebuchadrezzar came with an army against me, to deliver battle. Thereupon we joined battle; Ahuramazda bore me [*upastā*]; by the favor of Ahuramazda I smote that army of Nidintu-Bel exceedingly.... (Darius Behistun I:90-5)

It is interesting to note that in a number of these battles Darius' faction is not led by Darius himself but by a commander appointed by him. In those cases, however, it is Darius, rather than his lieutenant, who is the indirect object of the *upastā-* phrase.

Saith Darius the King: One man by name Ciçantakhma, a Sagartian — he became rebellious to me; thus he said to the people, "I am king in Sagartia, of the family of Cyaxares." Thereupon I sent off a Persian and Median army; a Mede by name Takhmaspada, my subject — him I made chief of them. Thus I said to them: "Go forth; the hostile army which shall not call itself mine, that do ye smite!" Thereupon Takhmaspada with the army went off; he joined battle with Ciçantakhma. Ahuramazda bore me [*upastā*]; by the favor of Ahuramazda my army smote that rebellious army and took Ciçantakhma prisoner.... (Darius Behistun II:78-88)

If *upastām bara-* means no more than 'assist', there is no reason that Takhmaspada and the other commanders should not be beneficiaries. Instead, Darius, though not even present, is the recipient of the *upastā-*.

I think we may find a necessary clue to the sense of *upastā-* if we direct our attention to the final occurrence of the word in the Behistun text.

This is in Darius' concluding statement, in which he sums up his victories and describes their significance.

Saith Darius the King: This which I did, in one and the same year by the favor of Ahuramazda I did; Ahuramazda bore me [*upastā*]-, and the other gods who are.

Saith Darius the King: For this reason Ahuramazda bore [*upastā*]-, and the other gods who are, because I was not hostile, I was not a Lie-follower, I was not a doer of wrong — neither I nor my family. According to righteousness I conducted myself. Neither to the weak nor to the powerful did I do wrong. The man who cooperated with my house, him I rewarded well; whoso did injury, him I punished well. (Darius Behistun IV:59-67)

Putting this in the context of Darius' situation, we can get the impression that the notion of *upastā*- is a very significant one indeed. The death of Cambyses, the last son of Cyrus, and the assassination of the pretender Gaumata left Cyrus' empire at the brink of dissolution. Darius thus found himself only one of a number of claimants to kingship, either over the empire as a whole or over one of its constituent nations. Under the circumstances, Darius' position was less than an ideal one. While he could, and indeed did, make a claim to a blood relationship to Cyrus by means of a rather complex genealogy, he was clearly, as Herodotus notes, at the time of Cambyses' Egyptian campaign "...a member of Cambyses' guard and not yet of any particular importance" (Herodotus 1984:261). Darius' edge over his competitors, in his view, comes from his piety: as a faithful adherent of Ahura Mazda, the legitimacy of his claim to kingship rests upon the support of the god.

I would like to suggest that this is where we can locate the true role of *upastām bara*- in the Behistun text. If we assign to *upastā*- the meaning 'foundation' or 'base' which we extracted from Ossetian and Middle Persian, we find that this phrase turns out to be comparable to the Akkadian *išdān rakāsu* and *išdān kunnu* 'make fast', 'establish the foundations', which likewise deal with sovereignty and authority in the political and military spheres.³ It is Ahura Mazda's intervention on behalf of Darius that elevates him from the crowd of usurpers and adventurers to the status of the legitimate heir of Cyrus.

Unto Ahuramazda thus was the desire: he chose me as (his) man in all the earth; he made me king in all the earth.

I worshipped Ahuramazda. Ahuramazda bore me [*upastā*]-.... (Darius Susa F 15-19)

Compare the accounts of Darius' campaigns described above with those in the fifth column of Behistun, added several years after the main body of the text. During the course of Darius' second and third years on the throne, the King suppresses another revolt in Elam and launches an invasion of Scythia, but the narrative deals with these episodes in a manner quite different from that of the earlier battles. Whereas the wars of Darius' accession year are characterized as conflicts among full-fledged sovereigns — "nineteen battles I fought; by the favor of Ahuramazda I smote them and took prisoner nine kings" (Darius Behistun IV:5-7) — in the later battles Darius' antagonists are identified not as 'kings' (*xšāyaθiya-*) but as 'chiefs' (*maθišta-*). Here also, rather strikingly, the familiar *Auramazdāmaiy upastām abara* is absent. By the interpretation we have been employing, the significance of the absence is clear: Darius' authority has been consolidated, and neither the Elamites nor the Scythians pose a challenge to his throne. The question of Ahura Mazda's support does not arise.

We may note, finally, that *upastām bara-* is not restricted to situations of crisis. The phrase also occurs in general exhortations for civil welfare and dynastic security.

Saith Darius the King: May Ahuramazda bear me [*upastā-*], with the gods of the royal house; and may Ahuramazda protect this country from a [hostile] army, from famine, and from the Lie! ... (Darius Persepolis C 12-20)

Saith Darius the King: Ahuramazda is mine, I am Ahuramazda's. I worshipped Ahuramazda; may Ahuramazda bear me [*upastā-*]. (Darius Susa K 3-5)

The Old Persian *upastā-* thus seems to revolve around securing and maintaining royal authority. If we bear this in mind, we may be able to make sense of some forms found in texts of later provenance. Herzfeld (1938:340), for example, in discussing Old Persian *upastā-* remarks that a comparable form *apastān* is familiar as part of a formula *'pst'n 'L yzt'n*, found on Sassanid seals. He settles on a translation 'faith in the Gods', to be compared to Arabic *tawakkal 'alā llāhi*. At the same time he cites a legend *pa tanaš apastān* on the seal of Queen Dēnak.

...The difference of the preposition is hardly without significance: one could think "in his (or her) person is (our) faith," but such legends are otherwise in the first person.... (Herzfeld 1938:340)

If we insert the modified interpretation of *upastā-*, however, we have the opportunity to read the Sassanid *apastān o yazdān*, like the Old Persian

upastām bara-, as a reflection of the ruler's wish to have his authority "founded" on divine patronage.

The Dēnak legend, on the other hand, remains obscure, but it is worth noting that Herzfeld has suggested that *pa tanaš* could be interpreted reflexively. While this is of little help to him as long as he is reading *apastān* as 'faith', it is possible that we can view this as meaning something on the order of 'self-founded', to be interpreted in the sense of 'auto-crat'.

In one of the most important Middle Persian texts, the Paikuli inscription of Šapur I, one of the less clear passages runs as follows:

W štry- [hwt'dyhy A]HRc 'pst'm [
[W hšt]r-hwtwpy BATRš [
"And for(?) the rule of the realm [there will] also hereafter
[be] reliance [on? ...]" (Humbach and Skjærvø 1983:32)

While one would not be justified in hazarding a translation of such a fragmentary text, we may nevertheless note that the term which has been rendered by 'reliance' bears a conspicuous similarity to the Old Iranian *upastā-*. It is worth noting, moreover, that the presence of *šahr-x'atāyēh* 'realm-domination' places the text securely in the area of political or dynastic authority, the area associated in Old Persian with the phrase *upastām bara-*.

We may conclude, therefore, with the observation that the rather minor emendation which has been suggested for *upastām bara-* ('lend support', rather than 'bring help') has led to some suspicions about a specific use of the phrase in the realm of politics. It is in this sense of *upastā-* that we may find a legitimate forebear to Ossetian *bästä*. The semantic development which we would assume in this case would actually be quite familiar. The shift from 'power, authority' to 'geographical area associated with the exercise of power' may be seen, for example, in the dual sense of English *dominion*, or in the Russian pair *vlast'*/*volost'*.

After all this, what may now be said about Scythian Βαστακας? As so often happens in the case of proper names, the absence of the restraint which context places on possible interpretations leaves the options as plentiful as the investigator's ingenuity and audacity will allow. Among the possible interpretations, enjoying varying degrees of plausibility, may be listed the following:

1. Βαστακας may indeed be comparable to Ossetian *bästä*, from **upastā-*; in this case it would probably be preferable to take the derivative in the 'zemljak' sense, or perhaps even posit a more direct reading 'associated with authority or power' (cf. Sanskrit *kṣatr-īya-*).

2. Βαστακάς may represent a contracted form of a more complex name, such as the name *Upastabar* found among the addresses of an Achaemenid Aramaic papyrus of the fifth century B.C. (*'pstbr pq[y]d' zy b'rbī h̄lš wmtlbš* 'Upastabar the officer who is at Arbel, Ḥalšu (?) and Māt-âl-Ubâš (?) ...') (Driver 1965:27-28)

3. Βαστακάς may indeed be a cognate to Indic *upas-tha-* 'lap'; cf. Sanskrit (Lexicographers) *upasthaka-* 'membrum virile'.

4. Βαστακάς may have at heart something totally unrelated. As an example, one could see in *basta-* the past participle of *band-* 'tie' (*basta-* < IE **bʰnd-tó-*; cf. Av., OP *basta-*, Ossetian *bast*). It might be comparable, therefore, to Persian *bande* 'servant', OP *baⁿdaka-* 'subject king'.

University of Chicago

Notes

¹It may be noted that not only is the derivative *upas-tha-* virtually without a cognate in the Iranian languages, but a cognate to the primary stem *upas-* seems to be attested only in the Avestan compound *upas-puθrya-* n. 'pregnancy'. Here *upas-* has the meaning 'womb' rather than 'lap', a fact which may well be of significance in taking Ossetian *bästā* as an Iranian counterpart to *upas-tha-*.

²The Old Persian texts and translations cited here and for the remainder of the paper are those found in Kent 1953:116-57.

³The Babylonian text of the Behistun inscription renders *Auramazdāmaiṣ upastām abara* by *^du-ri-me-iz-da-a i-se-da-an-nu* (and other spellings) (von Voigtlander 1978:13 *passim*). Early interpretations of this passage took *da-an-nu* as the adjective *dannu* 'strong', and *i-se* as 'my is-', *is-* being compared to *asū* 'physician'. Both the AHW and the CAD (s.v. *sēdu*) now see in this phrase a II-' verb, the standard preterite form of which would be **isēd-* (inf. **sēdu*), augmented by the ventive suffix *-am-* and ending with the first person suffix *-ni*. This verb is glossed 'help'; von Soden regards it as a borrowing from Aramaic *s-c-d*.

Another possibility which may be entertained, however, is that the verb is to be read **isēd-* (inf. **esēdu*). This would be a borrowing from Aramaic *y-s-d* 'found'. The imagery connecting founding and royal authority is quite familiar in Hebrew, as witness such lines as Isaiah 28:16.

*lkn kh 'mr 'dny yhw
hnn ysd bšywn 'bn 'bn bhn
pnt yqrt mwsd mwsd....*

"Therefore my Lord Yahweh has said, "Behold, I am founding in Zion a stone, a stone well-tested, the costly cornerstone of a foundation well-laid...." If such an interpretation of *i-se-da-an-nu* is valid, the parallel with the imagery proposed for *upastām abara* is obvious.

References

- Abaev, V. I. 1949. *Osetinskij jazyk i fol'klor I*. Moscow-Leningrad: Izdatel'stvo Akademii nauk SSSR.
- _____. 1958. *Istoriko-ètimologičeskij slovar' osetinskogo jazyka* (Volume I: A-K'). Moscow-Leningrad: Izdatel'stvo Akademii nauk SSSR.
- Driver, G. R. 1965. *Aramaic Documents of the Fifth Century B.C.* Oxford: Clarendon Press.
- Herodotus (trans. Aubrey de Sélincourt). 1984. *The Histories*. New York: Penguin Books.
- Herzfeld, Ernst. 1938. *Altpersische Inschriften*. Berlin: Verlag von Dietrich Reimer.
- Humbach, Helmut, and Prods O. Skjærvø. 1983. *The Sassanian Inscription of Paikuli. Part 3.1: Restored Text and Translation*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Kent, Roland G. 1961. *Old Persian: Grammar, Texts, Lexicon*. New Haven: American Oriental Society.
- CAD. = Reiner, Erica (editor-in-charge). 1984. *The Assyrian Dictionary*. (Volume 15: S). Chicago: The Oriental Institute.
- Vasmer, Max. 1923. *Untersuchungen über die ältesten Wohnsitze der Slaven. I. Die Iranier in Südrussland*. Veröffentlichungen des baltischen und slavischen Instituts an der Universität Leipzig, No. 3. Leipzig: Markert & Petters.
- AHW. = von Soden, Wolfram. 1972. *Akkadisches Handwörterbuch* (Band II: M-S). Wiesbaden: Otto Harrasowitz.
- von Voigtlander, Elizabeth N. 1978. *The Bisitun Inscription of Darius the Great: Babylonian Version*. Corpus Inscriptionum Iranicarum I:2. London: Lund Humphries.
- Zgusta, Ladislav. 1955. *Die Personennamen griechischer Städte der nördlichen Schwarzmeerküste*. Prague: Nakladatelství Československé akademie věd.

Syntactic Subject in Georgian

Kevin J. Tuite

Akaki Šanize (1963) and Arnold Čikobava (1968) have pointed out the problems that arise when notions of grammatical relations imported from monoperosonal Standard Average European languages (SAE) are imposed on a polypersonal language like Georgian. In SAE one noun phrase in the clause gives evidence of a special morphosyntactic prominence, in that the verb agrees with it – and no other – for person and number. The Kartvelian verb, by contrast, can agree with two, occasionally three, arguments. Hence, being cross-referenced by the verb does not in and of itself bespeak subjecthood in the SAE sense.

The notion of ‘subject’ as a linguistic category has been, of course, a topic of discussion for centuries. In this paper, I will focus on one aspect of this multifaceted concept, which I will term **syntactic subject**, abbreviated S_{syn} . Briefly put, the syntactic subject manifests formal and functional properties of intra- and interclausal scope, not shared with other types of argument. Within the clause it will manifest special agreement-governing properties and bind reflexive and reciprocal pronominals. Outside the clause it will function as **pivot**, i.e., play a key role in reference maintenance (Heath 1977, Foley and Van Valin 1984:ch.4). Some examples from English, a language with a prominent category of syntactic subject (S_{syn}) will illustrate.

(1a) Matthew¹ steps forward and \emptyset^1 embraces Lemmy Kilmister.

(b)* Matthew¹ steps forward and Lemmy embraces \emptyset^1 .

(c) Matthew¹ steps forward and \emptyset^1 is embraced by Lemmy.

Only the S_{syn} (underlined) governs verb agreement in English. In clause chaining, zero anaphora ordinarily only occurs between the S_{syn} s of the respective clauses; in (1c) zero anaphora between the agent of the first clause and the patient of the second comes about by means of passivization, which promotes the patient to S_{syn} status. Keenan (1976), Li and Thompson (1976), Foley and Van Valin (1984) and others have argued for the view that S_{syn} is neither a universally salient category nor an all-or-nothing one. Rather, the parceling out of agreement, binding, and pivothood properties can range from extreme asymmetry (and

therefore a prominent S_{syn}) to relatively even distribution (as in languages like Archi, claimed to have no S_{syn} (Kibrik 1979)).

Let's look at Georgian in the light of the preceding observations. Čikobava (1968:215-27) made the important proposal that grammatical relations in Georgian be viewed in terms of "coordination" (*koordinacia*) between arguments and the verb. He distinguished the following degrees of coordination in his description of Old Georgian:

(a) principal coordinate — the verb, which assigns case to one or two core arguments, (e.g., a transitive verb in the aorist assigns NAR case to its agent and NOM case to its patient).¹

(b) major coordinate — an argument with which the verb agrees in both person and number (e.g., an inclusive 1st person core argument, the NOM NP in Series 1 and 3, the NAR NP in Series 2).

(c) minor coordinate — an argument with which the verb agrees in person only (e.g., a 3rd person DAT NP).

(d) least coordinate — an argument with which the verb only agrees in number (e.g., a 3rd person NOM NP with a I Conjugation Series 2 verb).

Here are some Old Georgian examples. (p = person agreement, n = number agreement) (Examples from Čikobava (1968)).

(2a) $\overline{\text{man}} \overline{\text{še-krib-n-a}} \overline{\text{močape-n-i}}$ (he-NAR gather-I2 disciples-NOM)

'he gathered the disciples together.'

(b) $\overline{\text{Petre-s}} \overline{\text{klit̃-n-i}} \overline{\text{h-kon-an}}$ (Peter-DAT keys-NOM has-IV1)

'Peter has the keys.'

(c) $\overline{\text{da-gu-bad-n-a}} \overline{\text{man čuen}}$ (bear-I2 she-NAR us-incl)

'she gave birth to us.'

Aronson (1976) investigated the category of subject in Old Georgian on the basis of criteria similar to Čikobava's. One difference between the two approaches is that Aronson considered number agreement to be a particularly important sign of intraclausal prominence, especially for 3rd person arguments, which are the least likely to govern agreement (Čikobava 1968:271-79). In (3) are shown number and person agreement patterns for 3rd person NPs in Old and Modern Standard Georgian.

Note the change in inversion constructions (Series 3 and IV Conjugation).²

3. Verb agreement with 3rd person NPs

	I Conj			II Conj		III Conj		IV Conj	
SERIES 1-	NOM	DAT	DAT	NOM	DAT	NOM	DAT	NOM	DAT
OG	p/n	∅	p	p/n	p	p/n	p	p/n	p
SG	p/n	∅	p	p/n	p	p/n	p	p	p/n
SERIES 2-	NAR	NOM	DAT	NOM	DAT	NAR	DAT	NOM	DAT
OG	p/n	n	p	p/n	p	p/n	p	p/n	p
SG	p/n	∅	p	p/n	p	p/n	p	p	p/n
SERIES 3-	DAT	NOM	---	NOM	DAT	DAT	---	NOM	DAT
OG	p	p/n		p/n	p	p	---	p/n	p
SG	p/n	p		p/n	p	p/n		p	p/n

(p = person, n = number, ∅ = no agreement)

For Aronson, who was determining subjecthood on a series-by-series basis, calculating across person and lexical verb class, the Old Georgian Series 2 had no clear subject. In any event, it is clear that intraclausal prominence is the resultant of several independent factors: verb conjugation class, tense/mode series, argument role, NP type (1st/2nd vs. 3rd person), and the more subtle factors determining the choice of agreement patterns in the minimal pair shown in (4) (Tschenkéli 1958:488).

- (4a) es saxl-i da-∅-u-jd-a čems mšobleb-s bevri pul-i (this house-NOM cost-II2 my parents-DAT much money-NOM)

'This house cost my parents a lot of money.'

- (b) es saxl-i da-∅-u-jd-a-t čems mšobleb-s bevri pul-i

'My parents spent a lot of money on this house.'

Less has been said about other 'subjecthood' properties in Georgian. Harris (1981:ch.3) considers the binding of reflexives to be a key criterion of S_{syn} (cf. Chomsky 1981). In normative Standard Georgian (SG), those argument types governing number agreement in all persons are also the only ones to bind the reflexive *tavis* (*tavi*), as in (5).

- (5a) vano i-ṛmun-eb-s tavis tav-s
(Vano-NOM convince-I1 him-REFL-DAT)

- (b) vano-m da-i-rçmun-a tavisı tav-i.
(Vano-NAR convince-I2 him-REFL-NOM)
- (c) vano-s da-Ø-u-rçmun-eb-i-a tavisı tav-i.
(Vano-DAT convince-I3 him-REFL-NOM)

'Vano convinces/convinced/has convinced himself.'

As far as reference maintenance is concerned, Georgian does not have the tight control over zero anaphora found in English. SG allows "unemphatic pronoun drop" (Harris 1981:32-38), that is, any NAR, NOM, or DAT argument slot need not be filled by a surface NP. According to text counts performed by L. Enuşize (1978) zero anaphors are more likely to be found in "subject" (= S_{syn}) or "indirect object" position than in the "direct object" slot. On the basis of facts such as those discussed above, one can claim that SG has a category of syntactic subject, though not an exceptionally prominent one.

The surface coding patterns of the nonstandard Georgian dialects differ to varying degrees from the SG situation. Particularly noticeable are the differences between east Georgian dialects (e.g., Xevsurian, Pşavian, Moxevian) and western ones (e.g., Gurian, Imeretian, Aşarian). In a monumental study published in 1920, Şanize presented data from several dialects which indicate a fairly consistent east-west split with regard to the distribution of the verbal prefix *h-/s-/x-*, which marks a 2nd person "subject" or 3rd person "object." Many eastern dialects preserve this affix in some form, while almost all western dialects lack it, as exemplified below (Şanize 1920:ch.7).

(6a) XEVSUR: ıgi ga-s-ıtex-s mas man ma-x-par-a ıgi mas

(b) GURIAN: ıgi ga-ıtex-s mas man mo-par-a ıgi mas

(he-NOM break-I1 it-DAT) (he-NAR steal-I2 it-NOM him-DAT)
'he breaks it' 'he stole it from him'

While 1st/2nd person agreement patterns remain basically the same across the Georgian-speaking area, agreement with 3rd person arguments shows interesting variation. The Xevsur verb can agree for person with two such NPs whereas the Gurian verb agrees with only one.³ The number agreement situation is rather complicated, but also shows an east/west distinction. Şanize (1915), Čikobava (1968:276-77) and Kızıría (1974) cite examples of 3rd plural agreement with "non-subject" NPs from several dialects. Instances of agreement in *-n-* with plural NOM NPs in Series 2, as in Old Georgian, were collected by Şanize (1915:18-19) from Pşav and Xevsur speakers. Agreement in *-t* with plural DAT NPs is also

widely attested in eastern Georgia (Čikobava 1968:276-77). Some examples:

- (7a) (man) mta-n-i da-laq-n-a (he-NAR climb-I2 mountains-NOM)
'he climbed over the mountains.'
- (b) kac-ma $\overline{\emptyset}$ -u-txr-a-t mat (man-NAR said-I2 them-DAT)
'the man said it to them.'
- (c) igi $\overline{\emptyset}$ -u-qvir-i-t glexeb-s (he-NOM shout-III1 peasants-DAT)
'he shouts at the peasants.'

Some western Georgian dialects (Imeretian, Račan, Lečxumi) employ a morpheme *-qe* which cliticizes onto the verb to indicate plurality of a DAT or (direct object) NOM argument. Kiziria (1974) gives numerous examples; the following are from Imeretian.

- (8a) mi-čer-e-qe $\overline{\text{čeril-i}}$ bošeb-s (write-I2 letter-NOM children-DAT)
'You wrote the letter to the children.'
- (b) $\overline{\text{mat ak-qe}}$ isar-pšvild-i (they-DAT have-IV1 bow-arrow-NOM)
'They have bows and arrows.'
- (c) e bošveb-i da-v-zard-e-qe $\overline{\text{(this children-NOM raise-I2)}}$
'I raised these children.'

Both Pšav-Xevsur and Imeretian employ the plural DAT marker (*-t* or *-qe*) in inversion constructions with a plural "subject." (For an examination of number marking in literary Georgian inversion, see Aronson 1976, Sarjvelaze 1981). The western dialects Gurian and Ačarian only allow number agreement with one 3rd plural argument, as in SG. Another trend observed in the west involves the spread of the affix *-en* as a 3rd plural marker to screeves where SG employs *-es* or *-t*. Kiziria (1974:83-84) gives some Lower Imeretian examples:

- (9a) $\overline{\text{çkuereb-ma}}$ u-txr-en sulel-ş (clever-NAR said-I2 silly-DAT)
'The clever ones said it to the fool.'
- (b) simid-i šu-u-çam-i-en $\overline{\text{mat}}$ (corn-NOM eat-I3 they-DAT)
'They have eaten the corn.'

This is evidence of a growing uniformity of cross-reference marking for a particular class of arguments — those that correspond to the S_{syn} of Standard Georgian. The most extreme instance of this trend is represented by the Lower Ačarian subdialect spoken in southwest Georgia near the Turkish border. The NAR case, used to mark the agent of I and III Conj verbs in Series 2 in SG, has been extended to almost all tense/mode paradigms including Series 1, and all verb conjugation classes

(Boeder 1979:445, Harris 1985:15.3.4). In some instances the NAR case alternates with the NOM or DAT in the speakers' usage.

(10a) *turkeb-ma paṭron-ob-en sopel-i* (Turks-NAR rule-I1 village-NOM)
'The Turks rule the village.'

SG: *turkeb-i paṭron-ob-en sopel-s*

(b) *sakonel-ma mo-d-i-s* (cattle-NAR come-II1)

'The cattle are coming.'

SG: *sakonel-i mo-d-i-s*

(c) *kaṭa-m mo-xuc-d-a* (cat-NAR grow-old-II2)

'The cat grew old.'

SG: *kaṭa-∅ mo-xuc-d-a*

(d) *kaṭeb-ma sačečl-it de-e-čxliṭ-en saçqal kal-i*⁴ (fiends-NAR
comb-INSTR stick-I3 poor woman-NOM)

'The fiends stuck the poor woman with a carding comb.'

SG: *kaṭeb-s sačečl-it da-e-čxvliṭ-a-t saçqal kal-i*

There is little that can be said concerning binding phenomena in the dialects. Martirosovi (1964) has argued that the use of *tavisi* as a possessive adjective coreferent with the syntactic subject is a relatively recent innovation in Georgian (9th-11th centuries). An equivalent pronominal is not found in the other Kartvelian languages, and in the eastern dialects of Georgian there is frequent "misunderstanding" (*gaugebroba*) as to its use (Martirosovi 1964:118). The implication is that subject prominence is greatest in SG than in Old Georgian, and is especially weak in the east.

I have summarized much of the preceding discussion in the chart given in (11), where several dialects as well as Old Georgian and SG are compared for subject prominence. For each dialect, affixes that agree with 3rd person arguments are given. Each column represents an argument class; e.g., in Old Georgian, column A displays that argument class cross-referenced in the verb by *-a/o/s* (depending on the screeve) and by *-s/n* if plural, and which is marked in the NAR or NOM case (also dependant on screeve, though the case assignment pattern crosscuts the affix assignment pattern).

11. Marking of 3rd person NPs in Georgian⁵

	OLD GEORGIAN			PŠAV-XEVSUR		
	A	B	C	A	B	C
PERSON:	-a/o/s	h/s/∅	∅	-a/o/s	x/s/∅	∅
NUMBER:	-s/n	∅	-n-	-s/n	-t	-n-
CASE:	NAR,NOM	DAT	NOM	NAR,NOM	DAT	NOM
BINDING:	not until 11th C.			weak, if at all		

	STANDARD GEORGIAN		RAČAN		LOWER AČARIAN
	A	B	A	B	A
PERSON:	-a/o/s	h/s/∅	-a/o/s	∅	-a/o/s
NUMBER:	-s/n OR	-t	-s/n	-qe	-(e)n @
CASE:	NAR,NOM	DAT	NAR,NOM	DAT,NOM	NAR @
BINDING:	tavisi S _{syn} -bound		?		?

(@ = in alternation with other forms)

In Lower Ačarian one argument type plays a dominant role in governing verb agreement. Further, a set of agreement and case morphemes are coming to mark all and only appearances of this argument type, replacing the more complex, screeve-determined agreement patterns found in other dialects. *In fine*, we can say that Lower Ačarian has a well-developed category of syntactic subject. How prominent the S_{syn} is from the perspective of interclausal reference maintenance functions remains to be determined; one would expect it to play a more dominant role than, say, the S_{syn} of Standard Georgian. Like the latter, the Ačarian S_{syn} patterns in a nominative-accusative fashion. The situation in Old Georgian and Pšav-Xevisur is rather different, as (11) indicates. Three classes of arguments are cross-referenced in the verb in all three persons. No one of them exclusively binds the reflexive *tavisi*, in contrast to SG. As far as can be told, there is no S_{syn} in these dialects, or there is at best a weakly developed one. Rather than declare a category of subject by fiat (e.g., that NP marked in the NOM case, or the one corresponding to the S_{syn} in SG), the more straightforward approach, I would say, is to leave the category out of one's description of

Pšav-Xevsur morphosyntax entirely, until further investigation settles the issue.

University of Chicago

Notes

⁰It is with particular pleasure — and a bit of awe as well — that I dedicate this paper to Professor Aḡaḡi Šanize. There are few academic careers indeed which can match his in any single dimension, much less all three: length, breadth and depth.

¹An explanation of the terminology and abbreviations:

case: NOM = nominative (-i/Ø), NAR = narrative, also called ergative (-m), DAT = dative (-s). These cases are not formally distinguished for 1st and 2nd person pronominals.

conjugation: Georgian has four formally-characterized lexical verb classes. These are: I Conjugation (predominantly transitive), II Conjugation (usually intransitive, includes passives derived from I Conj verbs), III Conjugation (comprises atelic ‘middle’ or ‘medioactive’ verbs denoting motion, habitual behavior, emission of stimulus), IV Conjugation (inverse verbs, mainly denoting perception, emotion or sensation).

screeve (from Georgian *mčkrivi* “row”): a verb paradigm as usually found in a grammar or textbook — specified for tense, voice, aspect, etc., and only varying for person and number.

series: the three tense-mode series of SG (some dialects have a fourth, which will not be relevant here) are groups of screeves with common semantic and case-assigning characteristics. They are: Series 1 (present/future, imperfect, etc.), Series 2 (aorist, optative, etc.), Series 3 (present perfect, pluperfect, etc., often used to indicate evidentially-inferred past action).

In glosses of verbs, Roman numerals indicate conjugation class, Arabic numerals indicate series.

²This change has not quite reached completion in normative SG. In inversion constructions, a 1st or 2nd person patient (NOM), or 3rd person agent/experiencer (DAT) can be marked for plurality by -t. Should these occur together, -t can only mark the number of the former: *mo-v-s-çon-var-t* can only mean ‘he likes us’ or ‘they like us’. In colloquial usage, though, it can also mean “they like me” (Tschenkéli 1958:460, 510).

³I am not treating the “version vowels” which precede the verb root as cross-referencing affixes, even though they have strong associations with valence (e.g., verbs with -i- are usually monovalent; -e- implies the presence of an indirect object).

⁴Series 3 and IV Conj verbs with NAR case agent/experiencers are also attested in the Lower Kartlian villages Çerakvi and Sioni. Imnaišvili (1961) cites examples such as: *deda-m çqal-s mu-a-k-s* (mother-NAR water-DAT bring-IV1) ‘mother brings water’ (SG: *deda-s çqal-i moakvs*). This phenomenon may be due to the influence of Armenian- and Azeri-speaking communities which completely encircle these villages.

⁵A bit more should be said about the hierarchy of argument types in Georgian. Čikobava (1968:271-79) presented evidence that in Old Georgian 1st person NPs were more likely to govern number agreement than 2nd person NPs, which in turn were more prominent than 3rd person NPs. Among the latter, plural arguments in *-n*-governed number agreement more readily than plurals using the old collective desinence *-eb-*. In SG all 1st and 2nd person main arguments (i.e., arguments that, in the 3rd person, are in the NOM, NAR or DAT case) govern agreement for person and number. (Cooccurrence of 1st/2nd person direct and indirect objects, however, is not allowed in SG; the would-be direct object is demoted to oblique status (Harris 1981)). 3rd person animate arguments usually govern number agreement when functioning as S_{syn} in SG; inanimate NPs usually do not. *Ceteris paribus*, animate NPs are far more likely to show up as subjects (see Tschenkéli 1958:486-90 for some instructive examples). Especially intriguing are instances of focus-conditioned inversion, as in example (4b), which indicate that, in some cases at least, selection of S_{syn} is independent of the structure of the verb stem and its case-assigning properties. Further investigation into this should prove amply rewarding.

References

- Aronson, Howard. 1976. "Grammatical subject in Old Georgian." *Bedi Kartlisa* 34:220-31.
- Boeder, Winfried. 1979. "Ergative syntax and morphology in language change." *Ergativity*, Frans Plank, ed., New York: Academic Press: 435-80.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Čikobava, Arnold. 1968. *Marṭivi činadadebis problema kartulši, I*. Tbilisi: Mecniereba.
- Enukize, L. 1978. "Zmnis mier martul saxelta sixširuli daxasiateba kartulši," *Iberul-kaṽkasiuri enatmecniereba* 20:61-75.
- Foley, William and Robert Van Valin. 1984. *Functional syntax and universal grammar*. New York: Cambridge University Press.
- Harris, Alice. 1981. *Georgian syntax: a study in relational grammar*, New York: Cambridge University Press.
- _____. 1985. *Diachronic syntax: the Kartvelian case*. New York: Academic Press.
- Heath, Jeffrey. 1977. "Choctaw cases." *Berkeley Linguistics Society* 3:204-13.
- Keenan, Edward. 1976. "Towards a universal definition of subject." *Subject and topic*, Charles Li, ed. New York: Academic Press: 303-33.
- Kibrik, A. E. 1979. "Canonical ergativity and Daghestan languages." *Ergativity*, Frans Plank, ed. New York: Academic Press: 61-77.
- Imnaišvili, G. 1961. "Kvemokartluris zogierti tavisebureba." *Kartvelur enatas struṭṭuris saṭixebi* 2:150-55.
- Kiziria, A. 1974. "Kvemdebare-šemasmenlis urtiertoba kartuli enis dasavluri dialeṭebis mixedvit." *Kartvelur enata struṭṭuris saṭixebi* 4:75-91.
- Li, Charles and Sandra Thompson. 1976. "Subject and topic: a new typology of language." *Subject and topic*, C. Li, ed. New York: Academic Press: 456-89.

- Marčtirosovi, A. 1964. "Ukukcevit-kutvnilebiti nacvalsaxelesi kartvelur enebši." *Iberiul-kavkasiuri enatmecniereba* 14:109-23.
- Sarjvelaze, Zurab. 1981. "Zur Geschichte einer syntaktischen Erscheinung." *Georgica* 4:86-87.
- Šanize, Akaši. 1915. "Kartuli kiloebi mtaši." [Reprinted in Šanize 1957:9-19.]
- _____. 1920. "Subiekturi prepiksi meore pirisa da obiekturi prepiksi mesame pirisa kartul zmnebsi." [Reprinted in Šanize 1957:111-266.]
- _____. 1957. *Kartuli enis sstrukturisa da istoriis saqitxevi, I*. Tbilisi: TSU Gamomcemloba.
- _____. 1963. "Le sujet grammatical de quelques verbes intransitifs en géorgien." *BSLP* 58:1-27.
- Tschenkelí, Kita. 1958. *Einführung in die georgische Sprache, I*. Zürich: Amirani Verlag.

OTHER STUDIES

Gilyak Internal Reconstruction, 3: Ligneous Matter

Robert Austerlitz

Apparatus: gS and gA refer to the Saxalin and the Amur dialects of Gilyak (Russian *nivx*, gS *ñi-γvη*). Examples from gA are provided only when they are of interest in the argumentation; they are from Savel'eva and Taksami (1970) and can be controlled there. Data from gS are my own; see Austerlitz 1956, 1982, and 1984. Supporting examples are given in the cumulative table at the end and referred to by number (##). Plants were originally identified, from specimens and from descriptions, by the late Misao Tatewaki (Faculty of Agriculture, Hokkaidō University) and Hiromiti Kōno (Hokkaidō University of Education).

The typical Gilyak noun is monosyllabic, like the following, taken at random: *ñηαḡ* 'eye', *m̥la* 'ear', *ux* 'nose', *gōřqř* 'throat', *dot* 'arm', *ηacx* 'leg/foot', *coηqř* 'head', *duñm̥η* 'finger/toe'. As these examples show, the typical noun also contains rich initial and final clusters.

These clusters suggest earlier polysyllabicity and, what is more, polysemicity. With a little imagination it is not too daring to segment *ñηαḡ* 'eye' into **ñ-ηa-ḡ*, where **ñ-* may be connected with *ñu-* 'see' (also *nñu-* and *inty-* 'behold', cf. *ñetf* 'face'), *-ηa-* with a very large number of words with that same segment, all referring to animal life or living matter (cf. also *ηa* 'fur, skin', *ηa ηanγ-* 'go hunting'), and final *ḡ* with a number of nominalizing suffixes. The greater the complexity of the cluster, the greater the number of meaningful segments for which the method must account. The identification of the meaningful segments must be based on triangulation, pinpointing the recurrence of one and the same segment, in as many contexts as can be found, with, presumably, the same original meaning. This is the procedure which will be followed here and this accounts for the many examples adduced.

The word for 'tree, wood' is *c̥xar* in gS and *ciγr* in gA. We reconstruct **ci-ga-də*, with stress on **gá-* in proto-gS and with stress on **ci-* for proto-gA. When a reconstructed unstressed vowel cannot be specified it is written *ə*. Lenition of original intervocalic stops (the change to their corresponding fricatives), as in **ga- > *-γa-*, is regular. Since *a* attracts

adjacent postvelars (*q*, *g*, *ḡ*, *ḡ*, also called uvulars), the velars **g* and **ḡ* in **ci-gá-də* > **ci-ḡá-də* become **ḡ*. In the case of *gA*, the original **g* is not adjacent to an *a* and therefore does not change to a post-velar: proto-*gA* **ci-gə-də* > **ci-ḡə-də*. Loss of the unstressed reduced vowels yields *gS* **c-ḡa-r* and *gA* *ci-ḡ-r* (*-d-* lenites to *-r-*). The Saxalin dialect only tolerates initial clusters of fortis obstruents such as *px-*, *tf-*, *kř-* but not mixed fortis-lenis or lenis-fortis clusters (**pḡ-*, **tv-*, **kr-*; **bx-*, **df-*, **gř-*) or lenis-lenis clusters (**bḡ-*, **dv-*, **gr-*). This accounts for *gS* *cḡař*, with initial *cḡ-* from **cḡ-*. Similarly, *gS* only tolerates fortis (voiceless) fricatives in absolute final position; hence *gS* final *-ř* in *cḡař* (< **əḡə* **-adə*) as against *gA* *-r* in *ciḡr*.

We assume that this word is historically polymorphemic. The putative morphemes of the reconstructed form are **ci-* 'plant matter, ligneous matter', **-gV-* 'straight' or 'upright, vertical', and **-də*, nominalizer. The last of these is extremely frequent in the language as an agent morpheme and is still productive in the verbal paradigm as a verbal noun, '-ing'; it will not occupy us further, except to contrast it with the final *ḡ* of *cḡař* 'group of trees at the edge of the tundra; (roughly:) forest'.

We will posit the meaning of 'homogeneous assemblage of congealed or cohering particles; mass' for the final *-ḡ* in *cḡař*. Evidence (by triangulation) for this can be sought in *caḡ* 'wood shavings, chips, splinters', *caḡ* 'water', *waḡ* 'moss', *gA* *ḡaḡ* 'fish-roe, milt' (*gS* *ḡawk*), *gA* *coḡ* 'blood' (*gS* *cox*), *daḡ* 'epidemic', and perhaps even in *baḡ* 'rock, stone', *gA* *goḡ* 'stomach, belly, abdomen', and *joḡ* (*gS*) 'newly-wed young woman: (gA) younger brother's wife'.¹ Other, less convincing examples and blatant counterexamples are listed under ##1-4 in the table.

The first example listed in support of the idea that *-ḡ* means 'mass' is particularly significant: *caḡ* 'wood shavings'. If *caḡ* can indeed be analyzed as **-cḡ-gV-*, then it would mean 'mass of *ca-*' etymologically. Is this **ca-* in any way connected (*a*-grade, Austerlitz 1982) with the **ci* for which we posited the meaning 'ligneous matter' above? If so, the form would justify the meaning: wood shavings are a mass of ligneous matter. Another of the exhibits adduced in evidence above: *caḡ* 'water'. Compare *cacf* 'swamp', *cati* 'well', *caf&cav-* 'splash'.

What other evidence have we for the element **ci-/ca-*? There is, first of all, *ca* 'ligneous matter under the outermost bark' (synonym: *dikř*), as if made to order. Other words with similar initial segments are less convincing but worthy of mention: *caḡx* (also *caḡ*) 'tip or top of a plant', where the final *-x* may conceal the meaning 'pointed', cf. *ux* 'nose', *nux* 'needle', *ax* 'tip, front end'; *camx* 'tree stump'; *ceř* 'branch, bough, twig'; *ceřf* 'conifer needle'; and *comř* 'leaf' (synonym: *planq*). Initial clusters are even less reliable because lenis *c* becomes a fortis *c*.

Nevertheless, note *c̥xaj* 'large, stout, old pitch-tree or fir (? *Picea jezoensis*)', *c̥xō̃vi* 'polygonum convolvulus (an annual herb)', and *c̥xō* 'sap of trees'. The last of these triangulates in too many directions. Its initial *c* is reminiscent of *cox* 'blood' and *caṣ* 'water'; its fricative may be connected with the postvelar fricative in *cõṣ* 'tree rings'. Finally, *c̥nyř* 'grass' deserves mention, even though its meaning is hardly reconcilable with 'ligneous matter'.

The second element in the reconstructed form **c̥i-gV-də*, **ḡa* in proto-gS, is ludicrously close to the gS nursery word for 'tree, wood': *ḡaq*. The final *-q* in the nursery form is a playful-creative suffix typical of nursery words and poetic language (see Austerlitz 1956). Rather than serving as an atavistic catalyst, the nursery form is simply an instance of the reduction of initial clusters (unmarked *c̥xar̥* as against nursery *ḡa-*) and the substitution of the playful-creative suffix *-q* for the final *-ř* in the unmarked form. If the nursery form, then, is not of much help in the search of an etymon for **ḡa-*, the two following nouns are: gS *ḡar̥* 'back, spinal column' (gA *gydr* — the two forms suggest an earlier **gV-də-də*, where *V* is *a* or *y*) and *ḡalmř* 'board, plank', folkloristically distorted into *qalṇř* (gA *gylmr*). The etymon which suggests itself for **ḡa-* would be either 'straight' or 'straight and upright (vertical, erect)'. The 'vertical' component can be supported to some extent by the verb *gypr-* 'be in a standing, upright position'. The vowel alternation *a:y* is fairly systematic in the language; for *-p-* in *gypr-* see *j-op-u-* in note 1.

The five plant names (§§5-9) in the table offer little support for the etymon proposed. They are given for the sake of completeness. *ḡazř* 'mushroom' is strongly reminiscent of Ainu *karús* and probably a loan. The Gilyak are traditionally mycophobic.

Let us then conclude that the proto-form meant only 'tree', i.e., 'ligneous matter' + 'straight' and/or 'upright' + agent morpheme.² The name of the substance or material, wood, was carried by the first element, **c̥i-*, which, as we have seen, has a broad combinatory range. We will therefore provisionally propose that **c̥i* was a classifier and that Gilyak was a classifier language with a classifier slot for 'substance.'

Our search for evidence that **c̥i* was a classifier takes us to *c̥ṇaj*, a religious term with secular accretions: 'idol carved out of wood; carved image of a god or spirit; picture, likeness; (aggressive or unguarded speech:) guy, chap, fellow'. If we take as a point of departure the idea that the religious meaning is the original one and that this is an object carved out of wood, then the initial segment is our **c̥i*. The medial segment — if we assume that religious beings are conceived as animate — would be the **ṇa* encountered above, in connection with *ṇṇax̣* 'eye'. This

ŋa refers to living matter, human, animal, perhaps plant, and maybe even divine. Final *j* is ever so widespread (see ##10-43) and also seems to be associated with animateness or with situatedness in a place, cf. *iv-* 'be (animate)', *fī-* 'be in a place'. Is *čŋaj* 'idol' then possibly analyzable as **ci-ŋá-i* 'ligneous in makeup' + 'animate' + 'being'?

Elsewhere (Austerlitz 1986:189) I have suggested, in connection with the element **do* 'obese, thick in circumference', that if nails (for hammering) were originally made of wood, then *gS třofř* 'nail' (*gA crows*) may also conceal the element **ci-*. The reconstructed form would then be **ci-dó-bə-də*, with the parcels A 'ligneous' + B 'thick' + C 'locus' + D 'agent' and perhaps with the immediate constituents (A (B C)) D, or: 'D the filler/stopper of' + 'BC the thick place' + 'which A is made of wood'. For *do-* 'fat, thick' note also *gA doγs* 'hewn block'.

Columbia University

Cumulative Table

	<i>qač</i>	lance, spear		<i>baj</i>	Vicia
	<i>nač</i>	gA plank-bed		<i>qorŋi</i>	elder
	<i>tnoč</i>	rope	25	<i>tlorŋi</i>	willow
	<i>ŋoč</i>	unrendered fat		<i>hajŋi</i>	Equisetum
5	<i>ğap</i>	Prunus padus		<i>hojŋi</i>	a berry
	<i>ğaqf</i>	Cornus seucicum		<i>hewŋi</i>	alder
	<i>ğarq</i>	Fritillaria		<i>gujŋi</i>	willow
	<i>ğarŋfq</i>	hard fungus on birch	30	<i>ŋarŋi</i>	fir
	<i>ğarŋvyt</i>	Arctostaphylos		<i>cari</i>	Empetrum
10	<i>tlyγi</i>	lynx		<i>curi</i>	a berry
	<i>vorŋi</i>	egg of the <i>lavŋi</i>		<i>ŋawri</i>	a lichen
	<i>lavŋi</i>	kind of horsefly		<i>haγi</i>	kind of onion
	<i>corŋi</i>	fly	35	<i>walmi</i>	Iris setosa
	<i>dapi</i>	kind of butterfly		<i>wařpi</i>	a fern
15	<i>ifrulŋi</i>	small gnat, sandfly		<i>kumpi</i>	a perennial herb
	<i>laxi</i>	kind of salmon		<i>tefq</i>	maple
	<i>deŋi</i>	kind of salmon		<i>cymγi</i>	Ovalifolium
	<i>bani</i>	kind of salmon	40	<i>cčoxvi</i>	an annual herb
	<i>miri</i>	toad		<i>tlaŋi</i>	reindeer
20	<i>memčyi</i>	?Tachydromus		<i>uski</i>	price
	<i>havsqi</i>	shell, outer cover		<i>přyski</i>	sacrificial dog
	<i>ğoj</i>	larch			

Notes

¹The word for 'bride' is suspect because of its initial *j*-. Few native nouns begin with *j*. In verbs it is always the reflex of the third-person preposed pronoun, **i*-. If we sever the *j*- from *joǰ*, the remaining segment **oǰ* triangulates into the intransitive verb **oǰ*- 'gather', its corresponding transitive *j-oǰ*- 'gather (people) into a designated area', *j-op-u*- 'collect (trans.)', and *wo* 'settlement, village' (*w*- is from **u*-, reciprocal). In this way *joǰ* 'bride' could be analyzed as being from **j*- third-person object, **o* 'gather', and *ǰ* 'homogeneous mass', i.e., originally, 'group of people gathered into an assemblage (in order to effect/discuss transfer of the young woman from one clan to another).' See Austerlitz 1990.

²Jean-Luc and Françoise Rivierre (Centre National de la Recherche Scientifique, Ivry-sur-Seine) inform me that in Camuhi (Cèmuhi or Wagap, New Caledonia) *àcùut* 'tree' is decomposable into *ā*-, agent prefix, and *cùut* 'être debout' (personal communication, 1984).

References

- Austerlitz, R. 1956. "Gilyak Nursery Words." *Word* 12:260-79.
- _____. 1982. "Gilyak Internal Reconstruction, 1: Seven Etyma." In: Howard I. Aronson and Bill J. Darden (eds.), *Papers from the Second Conference on the Non-Slavic Languages of the U.S.S.R.*, University of Chicago, April 28-29, 1981. Columbus: Slavica, 81-88. (= *Folia Slavica* 5:81-88).
- _____. 1984. "Gilyak Internal Reconstruction, 2: Iron and Questions Related to Metallurgy." In: Howard I. Aronson (ed.), *Papers from the Third Conference on the Non-Slavic Languages of the U.S.S.R.*, University of Chicago, March 23-25, 1983. Columbus: Slavica, 38-48. (= *Folia Slavica* 7:38-48).
- _____. 1986. "Contrasting Fact with Fiction: The Common Denominator in Internal Reconstruction, with a Bibliography." In: Dieter Kastovsky and Aleksander Szwedek (eds.), *Linguistics Across Historical and Geographical Boundaries. In Honour of Jacek Fisiak on the Occasion of His Fiftieth Birthday*. Berlin: Mouton de Gruyter, vol. 1:183-192.
- _____. 1990. "Gilyak *joǰ* 'bride' and *wo* 'village'." *Ural-Altaische Jahrbücher* 62:105.
- Savel'eva, V. N., and Č. M. Taksami. 1970. *Nivxsko-russkij slovar'*. Moscow: Sovetskaja ěnciklopedija.

Moldavian Linguistic Realities⁰

Donald L. Dyer

The term "Moldavian," when used in reference to a language, is a concept which continues to confound most scholars of the world. For most, the term "Moldavian language" has little or no meaning. Soviet scholars, and scholars influenced by Soviet language policy, however (e.g., Šišmarev 1953:113-20, Korlètjanu 1953:183, Graur 1955:311, and Il'jašenko 1983:84-86), contend that such a language exists, that it is a full-fledged literary language of the Romance family of languages, and that it has a rich literary history. This "language," they assert, is spoken within, and along the periphery of, the Moldavian Soviet Socialist Republic (MSSR). Interestingly, they attach no significance to the linguistic proximity of this language to Daco-Romanian (known more commonly as just "Romanian"), also of the Romance family. Korlètjanu, for example, in two other works, claims that the Moldavian literary language: (a) had been established as an independent language by the year 271 A.D., at which time two groups of Latin speakers (future Daco-Romanians and Moldavians) chose separate and different migratory paths (into present-day Romanian Banat, and Romanian Bucovina, Moldova and Bessarabia, respectively) as they moved northward in the Balkans through Dacia (1983:14); (b) experienced language "stabilization" as a result of the 19th century advent of a number of world-class writers, the most notable among whom were Donič, Krjangè, Aleksandri and Eminescu (1966:530); and (c) is an independent literary language of the eastern branch of the Romance family, which is linguistically on a parity with the other Eastern Romance languages of Romanian, Aromanian, Megleno-Romanian, Istro-Romanian and the already extinct Dalmatian (1983:10). (Important to note here is that the majority of scholars, e.g., Marioțeanu et al. 1977:111, treats Aromanian, Megleno-Romanian, Istro-Romanian and Daco-Romanian (called "Romanian" by Korlètjanu, above) as major dialects of one, greater Romanian language (Romanian *dialectele a limbii române*). Thus, Korlètjanu in effect is elevating to the status of languages what most scholars consider to be merely the dialects of a long-established literary language (i.e., Daco-Romanian). The possible intent behind such a claim will be discussed later.) Moldavian, Korlètjanu

states further, stands contemporarily apart from Romanian (Daco-Romanian) due to the influence of the East Slavic languages (1983:18).

There is another "Moldavian," and we will briefly mention it at this juncture, as it will later enter into our discussion of that language spoken in the MSSR. We refer more specifically to a Moldavian dialect (Romanian *subdialect moldovenesc*) of Daco-Romanian. This dialect of Daco-Romanian, as described by non-Soviet and non-Soviet influenced scholars (e.g., Rusu 1984:208-40 and Marioțeanu et al. 1977:130-45), is spoken in the Moldova and Bucovina regions of Romania proper (in the districts of Suceava, Iași, Botoșani, Neamț, Bacău, Vaslui, Vrancea, and Galați), the MSSR and the areas of the Ukrainian Soviet Socialist Republic which border on the Socialist Republic of Romania. Said Moldavian Daco-Romanian has isoglosses which extend into northeastern Transylvania, northeastern Muntenia, and northern Dobrogea (Rusu 1984:208). In our later discussion of this particular dialect, we will point out individually the specific grammatical features which characterize it.

While it is universally accepted that a Moldavian dialect of Daco-Romanian exists (Soviet and Soviet influenced scholars denying, due to its inferential nomenclature, that it is spoken in the MSSR,¹ but accepting the idea that it is spoken in Romania proper), the status of a Moldavian language has been much argued. It is to this matter that we now wish to address ourselves in this article. First, however, we must limit the scope of our discussion by recognizing and eliminating improper avenues of analysis.

The status of Moldavian (or any other language of the world) as a "language" may be argued on three levels: sociopolitical, historical-cultural, and linguistic; the first two of which, in the sense that they are basically less objective concepts, remain outside the scope of the third, which is much more scientifically oriented.

The sociopolitical level of argumentation is, of course, the most abstract. If, for example, Soviet language policy decrees that Moldavian is an independent literary language, then, in the sociopolitical arena, the question of the existence of Moldavian becomes moot. In point of fact, we find such a claim manifested in several Soviet theses, the justification for the literary status of Moldavian being that it is the right of every group of people united by a common language to call the language they speak their own and to characterize it as unique:

...language becomes one of the principal characteristics of ... [a] social group, specific for [a] contemporary society which is called a nation.

[A] nation, as it is understood, is a collective of people who have a stable character and who were born as a result of an historical process. This collective uses a common language.... (Korlètjanu 1969:6.)

Does such a criterion exist that, if it were successfully established independent of other criteria, it could serve as the criterion for determining an ethnic entity? Yes, it exists. This basic, constant criterion is language.... [I]n a linguistic context a people is composed of all those who speak one language. (Korlètjanu 1983:4, citing Filin 1962:76, 852.)

The Moldavian national language is the language of the Moldavian socialist nation, the language of the Moldavian people from the Prut [River] to the left bank region of the Dnestr. (Čeban 1953:159.)

Since Leninist language theory is highly idealistic and hopelessly impregnable and immune to linguistic analysis, there is no point in, and we consequently will refrain from, arguing, sociopolitically, the status of the Moldavian language.

Less abstract, though still basically nonobjective, is the historical-cultural level of argumentation. Herein, though, lies the most substantive and authentic nonlinguistic ammunition in support of the independent status of a Moldavian language, and any typical piece of Soviet research on this matter will bare consistently what facts there are:

The last [of the early Romanians] went, in turn, along two courses: one — to the region of the Banat, to northeastern Transylvania, and from there to Bucovina, Moldavia and Bessarabia, and the other — to Wallachia and to the southeast of Transylvania.... [N]ot one wave embraces Muntenia (Wallachia) and Moldavia together. Moldavia and Muntenia do not have genetic and historical connections. (Sergievskij 1959:141-42, citing Philippide 1925-28:389-92.)

Effectually, there is no disputing a claim such as this either. However, establishing the ethnic Moldavians as a culturally and geographically separate group of people descended from earlier Common Romanians (separate from, say, the later Daco-Romanians) says of their language merely that it is spoken in another place, and, as particularly in the case of the Moldavians, under different political, social, and cultural circumstances.

This brings us to the task at hand — the linguistic status of Moldavian. We will argue the linguistic status of the Moldavian language, but not in the sense that we will take sides in the matter, i.e., either asserting or denying the existence of Moldavian as a separate and independent literary language. Our “argumentation” will lie more along the lines of a systematic comparison of the grammar of literary Moldavian, as presented in the first two volumes of *Limba moldovenjaskè literarè kontemporanè*, the Academy grammar of Moldavian (hereafter, volume I, *Leksikoložija*, being referred to as Korlètjanu 1969, and volume II, *Fonetika și morfoložija*, being referred to as Korlètjanu 1970),² with the grammar of Daco-Romanian, the language to which many scholars, some quite vehemently (Bruchis 1982, for example), have said Moldavian is strikingly similar, save for the difference in their orthographies. (Contemporary Moldavian is written in Cyrillic; Daco-Romanian, today, is written with Latin characters.) Thus, by making a clean, unbiased linguistic comparison of Moldavian and Daco-Romanian, we will be drawing scientific conclusions about the two languages, but will be leaving the task of formulating personal opinions about this language situation to the readers themselves, avoiding our own personal convictions, and, as a result, possible sociopolitical and/or historical-cultural entanglements.

According to the Soviet publications *Atlasul lingvistik moldovenesk* (1968) (hereafter referred to as ALM) and Korlètjanu (1966:528), Moldavian basically is spoken within the confines of the MSSR, though isoglosses of this Moldavian's dialects touch Romanian Bucovina, Moldova, and Dobrogea. Though it is difficult to find stated in print, there appear to be three dialects of Moldavian: Northern, Central, and Southern (Korlètjanu 1966:560).³ We are a bit surprised to read also that the literary language is “supradialectal.... [and] is superior to the dialects” (Korlètjanu 1969:11). We would expect this to mean simply that there is no one dialect or combination of dialects on which the literary language is based, and that the literary language in fact draws on a number of its local dialects for its own phonological, morphological, lexical, and even syntactic features. But later, we will come to realize that the true meaning of the term “superior,” as it is used here, is more on the order of “completely different from,” and that “supradialectal” actually more closely approximates “having *nothing* to do with,” the local dialects.

Daco-Romanian, on the other hand, is spoken in Romania proper (and, given the non-Soviet point of view, in the MSSR). It is generally considered to possess six regional dialects (i.e., Moldavian, Transylvanian, Maramureș, Crișan, Banatian, and Muntenian),⁴ and its literary language is based on the spoken language which surrounds the Romanian capital city of Bucharest, that is to say, the Muntenian subdialect (Romanian

grai) of the Muntenian dialect (Rusu 1984:3). There are three important subdialects of Muntenian Daco-Romanian: Muntenian, Oltenian, and Dobrogean, the first two sometimes being referred to jointly as "Wallachian" (Rusu 1984:201-2). (From here on, "Muntenian," unless otherwise specified, will refer to the Muntenian subdialect of the Muntenian dialect of Daco-Romanian.)

In order to focus our investigation, let us assume first that Moldavian and Daco-Romanian linguistically are two different languages. If this is to be so, then upon any basic examination of the two languages, it seems fair to say that substantial differences in their grammars should be notable. If and where the two grammars coincide, the most likely explanation of such instances would be the merger of certain common dialectal features found in both literary languages. After all, Moldavian of the MSSR, whether it is a language or a variant, or merely a dialect of Romanian (i.e., Daco-Romanian), is always characterized as being spoken in the same general geographic area as the Moldavian Daco-Romanian dialect of Romania proper.⁵ Even the most adamant supporter of the notion of a separate Moldavian language must accept that there is at least some dialectal, and, consequently, dialectal-feature overlap here. Weigand's linguistic atlas of Daco-Romanian (1909, specifically maps 33-66) actually shows that these two speeches form a dialectal continuum. Thus, in the ensuing examination of the two languages, it will be more constructive to look at four things instead of two: literary and dialectal Daco-Romanian, and literary and dialectal Moldavian.⁶

As representatives of dialectal Daco-Romanian, we will select for our study both the Muntenian and Moldavian dialects of Romania proper; Muntenian because it forms the basis for the literary language, and Moldavian because, as we have already pointed out, it is said to form a dialectal continuum with the language spoken in the MSSR, a situation we should examine more closely. We will consider the Muntenian dialect to be that dialect which is represented by the speech of the dialectal sites 417 [Afumătă] and 418 [Buftea] in Weigand's (1909) atlas, and the Moldavian dialect to be that dialect represented by site 635 [Dumbrăveni] in the same atlas.

For dialectal Moldavian, let us initially look at the speech of Kišinev, the capital of, and most centrally located city in, the MSSR. Though it would be possible to select a more minor dialect of Moldavian that would be closer in geographic proximity to the Moldavian Daco-Romanian dialect of Romania proper (and, thus, one that doubtless would prove to exhibit dialectal features common with Moldavian Daco-Romanian), this is not necessary — all central Moldavian dialects, from the western to the

eastern borders of the MSSR (a region which includes Kišinev), are characterized by the same dialectal features. Perhaps an even better reason for looking at the dialect of Kišinev, though, is that it is unarguably a major speech center, and a most likely candidate for the basis of a literary norm. For the purpose of this article, the Kišinev dialect of Moldavian will be that dialect represented by a conglomerate of the dialectal sites that surround Kišinev: 132 [Kožušna], 133 [Grètiešt], 139 [Čučulen'], 140 [Kolonica], 145 [Manojlešt'] and 148 [Jaloveni]. These are utilized by the aforementioned ALM.

Let us begin this examination of the two literary languages, and the three accompanying dialects, with a brief look at the reflexes found in each for several common lexica:⁷

Chart A

# Literary Daco- Romanian	Muntenian Daco- Romanian	Moldavian Daco- Romanian	Literary Moldavian	Kišinev Moldavian
1. zi 'day'	zi	zi	zi	zi
2. pîept 'chest'	(p)k'ept	k'ept	pîept	k'ept
3. fačem 'we do'	fačem	fašim	fačem	faš'im
4. meržem 'we go'	meržem	meržim	meržem	merž'im
5. vulpe 'wolf'	vulpe	xulpi	vulpe	xulpi
6. sęarę 'evening'	sęarę	sarę	sęarę	sari
7. kiņne 'dog'	kiņne	kine	kine	kini
8. luminę 'light'	luminę	lumniņę	luminę	lun'ini
9. bęjat 'boy'	bęjat	bęjat	bęjat	bęjet
10. bęrbat 'man'	bęrbat	barbat	bęrbat	bęrbat
11. žos 'down'	žos	žos	žos	ž'jos
12. semn 'sign'	semn	sęmn	semn	sęmn
13. visez 'I dream'	visez	zisęz	visez	ž'isęz
14. sę męargę 'he goes [subj.]'	sę męargę	sę margi	sę męargę	si margi
15. umblu 'I walk'	umblu	imblu	umblu	imblu
16. frate 'brother'	frate	frati	frate	frati

Examining these forms, it is immediately evident that something is amiss. First, and most strikingly, the Moldavian literary language forms appear

to coincide, to the point of virtual phonetic mimicry, with those of literary Daco-Romanian. Secondly, while, as expected, literary Daco-Romanian appears to incorporate in its own phonology the majority of characteristic dialectal features found in the Muntenian dialect, the basis for the norm, the Moldavian literary language in effect seems to be devoid of dialect-specific features of the Kišinev dialect. Finally (though in this case not so surprisingly, since we have already made the point), the similarity of reflexes for these lexical items, in Moldavian Daco-Romanian and Kišinev Moldavian, would tend to support the view of a dialectal continuum scenario for these two speeches. Let us now look at these first two points in more detail, taking the second of the two points first.

Closer examination of the phonologies of literary and Muntenian Daco-Romanian reveals that, in an overwhelming number of ways, the Muntenian dialect coincides with the literary language. Chart A, above, points to this fact, showing exact phonetic correspondences between literary and Muntenian Daco-Romanian in 15 of the 16 lexical items. Other sources (e.g., Rusu 1984:199-200) cite only three major divergences of Muntenian Daco-Romanian from the literary language (feature 1, below, being exhibited in the lexical item *piept* [#2 in Chart A]):

(1) palatalization of labials in various stages (cf. literary Daco-Romanian *piept*, but Muntenian Daco-Romanian (*p*)*k'ept*);

(2) soft pronunciation of the prepalatal fricatives *š* and *ž* as *š'* and *ž'* (e.g., *moaš'e* 'midwives,' *koaž'e* 'husks'); and

(3) changes in vowel quality in eastern Muntenia: *e*, *a*, *o* → *i*, *i*, *u*.

In stark contrast to the situation in literary and Muntenian Daco-Romanian, if we examine the lexical forms given for literary and Kišinev Moldavian, we find almost no correspondences at all (statistically, only one of 16 items, item #10, *barbat*, corresponds precisely). By comparing the reflexes given in Chart A, and extracting information from Sergievskij (1959:89-97), we are able to enumerate a considerable number of phonological differences between literary and Kišinev Moldavian. More specifically, we find in the Kišinev dialect (and, incidentally, in the Moldavian dialect of Daco-Romanian, as well — see the appropriate column on Chart A), but not in the literary language, that:

(1) after the consonants *s*, *z*, *š*, *ž*, *c*, and *z*, the vowels *e* and *i* become *a* and *i* (see *zi* [#1], *semn* [#12] and *visez* [#13] on Chart A);

(2) final, and medial *e* when pretonic, becomes *i* (see *kiine* [#7] and *frate* [#16]);

(3) etymological *i*, when initial and stressed, is preserved (see *umblu* [#15]);

(4) final *a* becomes *i* (see *sęarā* [#6], *luminā* [#8] and *sā męargā* [#14]);

(5) in initial position, stressed *a* tends to become *a* (e.g., *apā* 'water,' but literary Moldavian *apā*);

(6) the diphthong *ia* becomes *ie* when stressed (see *bajiat* [#9]);

(7) after the consonants *s*, *z*, *š*, *ž*, *c*, *z*, *m*, and *b*, the diphthong *ea* becomes the monophthong *a* (in final position, however, it becomes *e*) (see *sęarā* [#6] and *sā męargā* [#14]);

(8) the labials *p* and *b* are palatalized to *k'* and *g'* (see *piept* [#2]);

(9) initial *v* becomes *x* before a nonfront vowel (see *vulpe* [#5]);

(10) the labio-dentals *f* and *v* become the fricatives *š* and *ž* (see *visez* [#13]);

(11) original Latin or substratum *z*, and *ž* from Latin *i* (etymologically an initial consonant followed by *o* or *u*) are preserved (see *zi* [#1] and *žos* [#11]); and

(12) the affricates *č* and *ž* become the fricatives *š* and *ž* (see *fačem* [#3] and *meržem* [#4]).

If the speech of Kišinev is the dialect on which the literary language is based, it is apparent from the above twelve observations that the literary language bears little resemblance to that dialect. Though the discovery of the Moldavian literary norm among its local dialects would, for us, be at best merely a fringe benefit of this study, it would, nonetheless, be satisfying to find just some of the features of the literary language there; from what we have seen so far, the Kišinev dialect does not begin to "fill the shoes" of the Moldavian standard (again, only one of our 16 lexical items on Chart A corresponds precisely).

Perhaps we can find better phonetic correspondences and greater resemblance to the literary language among other of the local dialects. Since we already have seen what there is of the Central Moldavian dialect, let us now take a look at the same sixteen lexical items given in Chart A, above, as they occur in the Northern and Southern dialects of Moldavian, the forms, as we present them, being juxtaposed to the reflexes previously given for the Central (Kišinev) dialect and the literary standard. For the Northern dialect, we will examine the dialect represented by the speech of Okju-Alb (site 45 in the ALM), and for the Southern dialect, the dialect represented by the speech of Koštangalija (site 196 in the ALM). Both are centrally located sites in their respective dialect regions.

Chart B

#	Literary Moldavian	Central Moldavian	Northern Moldavian	Southern Moldavian
1.	zi	zî	zî	zî
2.	piept	k'ept	k'ept	k'ept
3.	fačem	faš'im	feš'em	faš'im
4.	meržem	merž'im	merž'em	merž'im
5.	vulpe	xulpi	xulpi	xulpi
6.	sęarę	sari	sari	sari
7.	kine	kini	kini	kini
8.	luminę	lun'ini	lun'ini	lun'ini
9.	bęjat	bęjet	bęjet	bęjet
10.	bęrbat	bęrbat	barbat	bęrbat
11.	žos	ž'jos	ž'jos	ž'jos
12.	semn	sęmn	sęmn	sęmn
13.	visez	ž'isęz	ž'isęz	ž'isęz
14.	sę męargę	si margi	si męargi	si męargi
15.	umblu	imblu ⁸	imbl	umblu
16.	frate	frati	frati	frati

Disappointingly, nor do these dialects seem to be the basis for the Moldavian literary language. In particular, where the Kišinev dialect differed substantially from the literary language in twelve areas of the phonology, the Southern (Koštangalija) dialect likewise differs in twelve, and the Northern (Okju-Alb) dialect differs in thirteen. So what does this leave us? Perhaps we have here a literary language which is based on a conglomerate of its local dialects as, say, the literary Macedonian language is based on the four local dialects of Veles, Prilep, Kičevo, and Bitola (Lunt 1952:5). A quick glance back at our various Moldavian dialectal reflexes (Chart B) shows us, however, that *nowhere* does the literary language coincide with the features displayed in the dialects — neither within any one dialect, nor with any combination of reflexes from different dialects. Consequently, the conglomerate theory also must be rejected. This leads us to the realization, by process of elimination, that literary Moldavian is a language which essentially has no model for speech among its dialects.⁹ It is here that we find the intent behind the claim that literary Moldavian is “superior” to its dialects and

“supradialectal” (Korlètjanu 1969:11). Such a statement, unless challenged, may be used to thwart any search for the literary norm of the language where it should naturally be found – somewhere in the dialects – and to ward off claims that the true speech of the MSSR does not, in fact, serve as the norm for the literary language. To this point, we have suppressed any urge to look directly to Daco-Romanian (particularly Muntenian Daco-Romanian) for an answer to this puzzling problem. Let us now turn back to Muntenian Daco-Romanian.

If we compare literary forms from Moldavian and Muntenian Daco-Romanian (as, for example, those on Chart A), we observe that wherever literary Moldavian is divergent from its dialects (and this is almost everywhere), it coincides quite nicely with Muntenian Daco-Romanian (which in all but one of our 16 cited lexical forms matches feature for feature the forms in literary Daco-Romanian). The only exception results from the different treatment of the palatalization of labials (see, again, *piept* [#2]). Having paid to this point particular attention to the merger of phonological features between literary Daco-Romanian and Moldavian, perhaps we should now discuss where the languages actually differ. A look at the phonologies of Daco-Romanian (as presented in the Academy grammar of Romanian, *Gramatica limbii române, I, Fonetică și morfologie* [1963]) and Moldavian (as given in Korlètjanu [1970]) reveals only one difference.

The diphthong *iɨ* in Daco-Romanian (DR) when found before a nasal consonant is a monophthong in Moldavian (M). We see this on Chart A in DR *kiine*, but M *kine*. Other examples of this type are DR *piine* ‘bread,’ but M *pine* and DR *miine* ‘tomorrow,’ but M *mine*.

Further, if we leave phonology and look at the morphologies of Moldavian and Daco-Romanian, we find other pan-dialectal features of Moldavian which are not present in the literary norm, and we note, again, only one divergence of literary Moldavian from the literary standard of Daco-Romanian. (With regard to morphology, we no longer will cite dialectal forms from Northern and Southern Moldavian. We will concentrate solely on Central (Kišinev) Moldavian, since in their morphologies all Moldavian dialects exhibit the same features.)

First, with regard to how literary and dialectal Moldavian differ, we find in the Kišinev dialect, but not in the literary language:

(1) a genitive/dative singular inflectional ending for masculine substantives in *-lu*, e.g. *omulu* ‘man [G/D]’ (cf. literary Moldavian *omuluj*);

(2) a preference for a vocative form of the feminine noun in -è, e.g., *Anè* 'Ann [voc.]' (cf. *Ana*);

(3) an invariable (for number and gender) possessive-genitive article in *a* (cf. *a, al, aj, ale*);

(4) dative personal pronoun forms in *mja* 'my,' *cja* 'your' and *luja* 'his,' as well as third-person genitive forms in *a luja* 'his' and *a lora* 'their' (cf. *mie, cie, luj, a luj* and *a lor*);

(5) demonstrative pronouns in *ajasta, aista*, etc. 'this/these' (cf. *astja, asta*, etc.);

(6) the loss of the simple perfect verb tense (which has been replaced by the compound perfect), e.g., *am vèzut* 'I saw' (cf. *vèzuj*);

(7) an invariable (for number and gender) auxiliary for the compound perfect in *o*, e.g., *o fèkut* 'he/she/it/they has/have done' (cf. *a fèkut* and *au fèkut*);

(8) the loss of the suffix -rè in the plural form of the pluperfect, e.g., *vèzuse* 'they had seen' (cf. *vèzuserè*);

(9) the preservation of *a* in the first-person plural of the present tense, making the form homonymous with the imperfect, e.g., *dam* 'we give,' *lukram* 'we work' (cf. *dèm, lukrèm*);

(10) retention of certain archaic and analogical forms of the present subjunctive as *sè xibè* 'he has [subj.],' *sè skribè* 'he writes [subj.]' (cf. *sè ajbè, sè skrija*);

(11) forms for the third-person plural of the present indicative in *rèmyu* 'they remain' (cf. *rèmyn*);

(12) an imperative form for *a veni* 'come' in *vinè* 'come!' (cf. *vine*); and

(13) the forms *amu* and *amuš* for the adverb 'now' (cf. *akum(a)*). (Rusu 1984:236-37.)

Secondly, concerning differences between literary Daco-Romanian and Moldavian, as was the case with the phonologies, in the morphologies, too, we observe only one item. The Daco-Romanian third-person pronoun *dînsul* 'he' (inclusive of morphologically varying forms for other genders and numbers, e.g., *dînsa* 'she,' *dînşii* 'they,' etc.) has a much wider use in Moldavian than it does in Daco-Romanian. In Daco-Romanian, for most speakers,¹⁰ *dînsul* functions as a "marked" pronoun which expresses a greater degree of politeness than the unmarked *el* 'he' (again, inclusive of all other morphologically varying forms), having much the sense of 'his lordship' or 'his worship.' In Moldavian, however, *dînsul* functions as the unmarked third-person pronoun, alongside *el* (particularly in constructions which require an accusative object after the prepositional direct object marker *pe* 'on,' e.g., *pe dîns*), having achieved at least equal status with, and in some areas eclipsing the use of, *el* (Korlètjanu 1970:239-40 and personal

communication with Dorin Urișescu). In this way, not only is *dynsul* used for the expression of nonpoliteness (*el*, normally), but due, consequently, to its extended role in Moldavian, it comes to function as the pronoun for inanimate objects. Such usage is not found in literary Daco-Romanian:

L'am vèzut pe romanul. Dynsul este pe masè.

'I saw the novel. It is on the table.'

(Personal communication with Dorin Urișescu.)

The syntaxes of literary Daco-Romanian and Moldavian both basically exhibit an S-V-O-type structure of word order in the simple sentence and clausal constructions. In nearly every area of syntax the two languages coincide. The only significant divergence seems to exist in the choice of complement to the verbal bases and complement-requiring constructions which characterize the two languages. Here, there are differences in the two syntaxes, but within the individual languages themselves, the choice of complement may vary depending on style or subdialect, factors which make it too difficult for us to compare and contrast these constructions in this study. See, however, Vulpe (1963), Diaconescu (1977) and Dyer (1985) for an explanation of these matters.

Rounding out our comparative survey of the grammars of literary Moldavian and Daco-Romanian, we now turn to their lexical components. Unquestionably, it is in this area that we find the greatest linguistic difference between these two languages, and the number of individual lexical items which differ between the two languages is considerable. Supporters of literary Moldavian point consistently to the same areas of the Moldavian lexicon as being unique, and different from Daco-Romanian. Take Bugadov (1953:127), for example. Reiterating material in Iordan (1921:187), he lists a typical series of 15 lexical items found in Moldavian which differ, he claims, from the same of Daco-Romanian: *bortè* 'hole' (cf. DR *gaură*); *čobotè* 'boot' (cf. *cizmă*); *koromyslè* 'yoke' (cf. *cobilîță*); *čolan* 'bone' (cf. *os*); *xulub* 'pigeon' (cf. *porumbel*); *carmarok* 'fair, market' (cf. *bîlcu*); *mycè* 'cat' (cf. *pisică*); *ogradè* 'courtyard' (cf. *curte*); *plešuv* 'bald' (cf. *chel*); *skrob* 'battered eggs' (cf. *jumări*); *prostire* 'sheet' (cf. *cearșaf*); *perli* 'plum tree' (cf. *prun*); *okripkè* 'violin' (cf. *vioară*); *omèt* 'snow' (cf. *zăpadă*); and *geb* 'bunch' (cf. *cocoasă*). What is not made clear, is that all the Moldavian forms are also listed in the Academy dictionary of Daco-Romanian (Macrea 1958). In all fairness, though, all but *mycè*, *plešuv* and *geb* are characterized as *regional* 'regional.' And lest we should think that mention of these "regional" forms is restricted to official, and, thus, more authoritative (here, Romanian Academy) sources, we need only look at a much less unabridged work, such as Schönkron 1952, to find more than half (in this

particular case, nine) of these items. (He does not include *koromyslè*, *carmarok*, *skrob*, *prostire*, *perli* and *okripkè*.) Of those which he cites, none is described as "regional."

Elsewhere, Korlètjanu (1983:18) lists as being specific to Moldavian *buxaj* 'bull' (cf. DR *taur*),¹¹ again *xulub*, and a number of toponyms found in the MSSR: *Solonec*, *Volotna*, *Xorodka*, *Xorodište*, *Slatina*, *Gradište*, *Vyškèuc*, *Kolikèuc*, *Temelèuc* and *Čepelèuc*.

Sergievskij's treatment of Moldavian-specific lexica (1959:177-82), is, however, the most extensive of the lot. He gives (pp. 177-79) "a series of [86] words, common to all of Moldavia and completely unknown in Wallachia," which includes, in one phonetic variant or another, thirteen of the fifteen items listed by Bugadov, above (not given are *okripkè* and *carmarok*); a list of thirteen Moldavian lexical items which are also found in Wallachia, but which "differ significantly phonetically" from the Daco-Romanian forms (e.g., Moldavian *bondar* 'bumblebee,' *bukšè* 'boxwood' and *mejdan* 'market square,' but Wallachian (Daco-Romanian) *bombar*, *bucea* and *maidan*, respectively) (p. 179); a group of forty-five words found in Moldavian (pp. 179-80) which "do not have Wallachian parallels" (e.g., *dubas* 'boat,' *filè* 'sheet of paper,' and *xojma* 'continually'); and, finally, relying on material taken from Tiktin's Romanian-German dictionary (volume 1, 1903; only entries from the first three letters of the alphabet), ninety-seven lexical items which are noted as Moldavian (e.g., *bandraburkè* 'potato,' *bènoš* 'rich' and *košarkè* 'basket') (pp. 180-82).

A second group of lexical items which are often (and, in our opinion, more justifiedly) cited as being unique to Moldavian includes numerous more recent borrowings from Russian, the bulk of which has not been borrowed into Daco-Romanian. Bugadov (1953:128) gives eighteen such words, among which are *klub* 'club,' *kino* 'cinema,' and *kolxoz* 'collective farm.' But it is Korlètjanu (1983) who is exhaustive. He devotes the better part of his book (pp. 47-75) to such lexical items. In fact, in this book (which would appear from its title, *The Moldavian Language Today*, to be a basic discussion of the state of the Moldavian language today and a sketch of its grammar), he virtually ignores all other aspects of Moldavian grammar (the first 46 pages of the book are devoted to the history of Moldavian and the culture of the people of Moldavia), as he treats in very large, separate chapters scientific-technical terminology (pp. 47-57), agricultural terminology (pp. 57-66), and sociopolitical terminology (pp. 66-75). It becomes clear from these findings that an examination of the Moldavian lexicon has been, and will continue to remain, basic to any discussion of the autonomy of the Moldavian language.

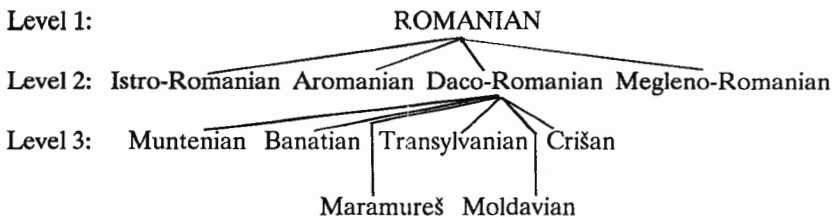
At this point, it is practical to summarize what we have learned with regard to the status of a Moldavian language, and to Moldavian as it compares with literary Daco-Romanian.

First, we have shown that the proclamation of literary Moldavian's separateness and independence as a language cannot be refuted on the political, historical and cultural levels of argumentation. Soviet language policy and Leninist ideology dictate the autonomy of the Moldavian people and, consequently, their language, and historical sources reveal that the Moldavian people did, in fact, migrate away from the other speakers of then Latin, later, Common Romanian. Linguistically, however, the task of separating literary Moldavian and Daco-Romanian into two distinct and independent languages is an endeavor which has as its reward a considerable amount of frustration. Not only is literary Moldavian not based on any one local dialect or combination of local dialects spoken in the Moldavian Soviet Socialist Republic (a fact which points to serious flaws and inadequacies in the actual codification process of literary Moldavian), but the literary language appears to be a virtual copy of literary Daco-Romanian. The exceptions are one phonological feature (before a nasal consonant, the Daco-Romanian diphthong *i̯* corresponds to the monophthong *i* in Moldavian), one morphological feature (the pronoun *dînsul*, which is often used to assign a greater degree of politeness in Daco-Romanian, has a much wider, unmarked use in Moldavian), and a certain (albeit large) set of lexical items. In this respect, the Cyrillic alphabet (in which Moldavian is written) serves as a cosmetic mask that blinds us to these otherwise obvious facts.

To be more specific, we have shown that literary Moldavian is not based on the dialect of Kişinev (a representative of the Central Moldavian dialect and, undoubtedly, the most likely candidate for a literary norm), the Northern dialect, the Southern dialect, or, again, on any combination of these or any more minor subdialects. Further, the three regional Moldavian dialects still form today a continuum with the Romania-proper portion of the Daco-Romanian dialect which non-Soviet influenced scholars call Moldavian Daco-Romanian, and exhibit the same phonological and morphological features as that dialect does. There is then no logical reason to cleave in two the speech of this general geographic area and to call what is spoken on the east bank of the Prut River a different language from that which is spoken on its west bank.

As a final note on the matter of Moldavian subdialects and dialects, here, we might venture a hypothesis as to why (as we mentioned earlier), Korlêtjanu (1983:10) has elevated to the status of languages what scholars overwhelmingly and quite legitimately consider to be merely

major dialects of Romanian. For most (Rusu 1984; Marioțeanu et al. 1977; etc.), the dialectal division of Romanian is as follows:



Seeing this, it becomes clear why Korlètjanu has made the claim that Moldavian, Romanian (for us, Daco-Romanian), Aromanian, Istro-Romanian, and Megleno-Romanian are all Romance languages, each of equal standing to the others. To raise what is in effect a “dialect of a dialect” (here, Moldavian Daco-Romanian), from Level 3, to the status of a language (Level 1), requires simultaneously the raising of everything in between (which in this case is all of Level 2). To call Moldavian a literary language while not calling, say, Istro-Romanian one, as well, would be tactical folly.

If literary Moldavian (again, as presented in Korlètjanu 1969 and 1970) has been standardized to stand on its own two feet as an independent language, separate from literary Daco-Romanian, then the process used for this standardization has produced little more than a linguistic house of cards. While the ideological theory used to support the status of a Moldavian language cannot be argued, the linguistic support for its separate and independent existence is tenuous at best. Linguistically, one different phonological feature, one different morphological feature and a group of different lexical items do not seem to be enough to divorce literary Moldavian from literary Daco-Romanian.

It is interesting that, in the past, Soviet scholars have pointed overtly to phonological and morphological features of the Moldavian variant of Romanian which clearly set it apart from Daco-Romanian (e.g., Sergievskij 1959:88-101), but have then gone on (in the matter at hand, only ten years later) to produce an Academy grammar of Moldavian (Korlètjanu 1969 and 1970) which in no way tolerates these same, important differentiating features in the literary language. Inclusion and standardization of just some of these dialect-specific features legitimately could justify the separation of the speech of the Moldavian Soviet Socialist Republic (i.e., Moldavian) from the rest of Romanian. Such incorporation, however, has not yet even been attempted, and now that Moldavian has become a “literary” language by virtue, if nothing else, of

the printed matter and studies devoted to it, it remains to be seen if it ever will.

University of Chicago

Notes

⁰Researching Moldavian has required a knowledge of Slavic, Romanian and dialectology, and I wish to thank those persons who were responsible for helping to develop my interests in those areas: Professors Victor A. Friedman (University of North Carolina at Chapel Hill), Howard I. Aronson (University of Chicago), Kostas Kazazis (University of Chicago) and Dorin Urişescu (Universitatea din Timişoara). I also would like to thank Professor Zbigniew Gołąb of the University of Chicago for first presenting me with the concept of "languages in contact in the Balkans," and Professor Aronson for suggesting that I read a paper on Moldavian at the conference of which this paper was a part.

¹To say that any form of Romanian (here, Moldavian Daco-Romanian) is spoken within the MSSR would be undesirable from the Soviet point of view. This would run contrary to the notion of complete separation of literary Moldavian and Daco-Romanian. It would also bring to the forefront discussion as to whether a trans-border dialectal continuum of Romanian exists, which consists primarily of the dialect of Moldova Romania and the speech of the MSSR.

²Within the last 35 years, various other language-standardizing works on Moldavian have appeared, which have helped to create the illusion that conducting research on this language is quite fashionable in Soviet linguistic circles, and that support for the independent literary status of Moldavian is spreading. The most notable of these works include: an orthographic dictionary (*Dikcionar ortografik* 1978); a literary journal (*Limba şi literatura moldovenjaskè* 1954-present); Moldavian-Russian and Russian-Moldavian dictionaries (Boršč, Podiko, and Solov'ev 1961 and Boršč, Korlètjanu, and Russev 1954, respectively); a two-volume dialectal atlas (*Atlasul lingvistik moldovenesk* 1968); and a Soviet bibliography of works published on Moldavian (Talmackaja 1978). Also, Korlètjanu (1983:83) makes reference to a new, 1982 edition of the Academy grammar of Moldavian, unseen by us.

³Sergievs kij (1959:88) presents a different dialectal division of the Moldavian within the MSSR. For him, Moldavian is better divided into two dialect regions: (1) the area to the north and east of the Jagorlyk and Kodymy Rivers, inclusive of the districts of Kamenskij and Rybnickij on the shores of the Dnestr, Krasno-Oknjanskij, Krutjanskij, and Birzul'skij; and (2) the territory to the south of the Jagorlyk River — the districts of Dubossarskij, Grigoriopol'skij, Tiraspol'skij, Slobodzejskij, and Anan'evskij. He also states (p. 92) that the local Moldavian dialects of the west bank of the Dnestr share common dialectal features with the Romania-proper dialects of Daco-Romanian spoken in Bessarabia, Bucovina, northern Moldova, and, to some extent, Transylvania.

⁴This, Rusu's breakdown (1984), which is the same as that given in Marioțeanu et al. (1977), is the most widely accepted, though certainly not the only classification of Daco-Romanian dialects. As in Rusu's work, special treatment is often given to the more specific local subdialects of the Transylvanian dialect (pp. 354-90), to the interesting subdialect of Țara Oașului (pp. 390-98) and to the subdialects of Daco-Romanian spoken outside the Socialist Republic of Romania: in Bulgaria, pp. 399-401; in Yugoslavia, pp. 401-3; in the U.S.S.R. (exclusive of the MSSR), pp. 403-5; in Hungary, pp. 405-7; and in the United States, pp. 407-20. In other descriptions of Daco-Romanian dialects (e.g., Neiescu 1971, Gheție 1964, and Rusu 1968), one dialect, particularly that of Transylvania, or more, may be merged with another to decrease the total number of dialects. This practice is taken to the extreme by Vasiliu (1968:152-93), who merges, primarily on phonetic criteria, all six Daco-Romanian dialects into just two: Muntenian and Moldavian.

⁵According to Soviet and Soviet-influenced sources, the Moldavian language is geographically contiguous to Moldavian Daco-Romanian (the Moldavian language in the MSSR, and the Moldavian dialect of Daco-Romanian in the Moldova region of Romania proper), whereas other sources consider the Romanian spoken in Moldova Romania and the language of the MSSR to be a dialectal continuum called "Moldavian Daco-Romanian." Rusu (1984:403), a contemporary Romanian linguist, makes it quite clear, from the following excerpt, what his position is with regard to the Moldavian-or-Moldavian-Daco-Romanian situation: "[f]or the local *Romanian* subdialects spoken ... on the territory of the U.S.S.R.[...] most of the particularities are common with the characteristics of the *local subdialects spoken in the Moldavian Soviet Socialist Republic...*" (emphasis added).

⁶We do not subscribe to the notion that literary Moldavian can be divorced completely from its local dialects. Thus, we will deal continuously with "dialectal" Moldavian as it relates to the literary language.

⁷In Chart A, the reflexes given for literary Daco-Romanian are drawn from Macrea (1958); those for Muntenian Daco-Romanian and Moldavian Daco-Romanian are gleaned from Rusu (1984) and Sergievskij (1959), respectively; and those for literary Moldavian and Kišinev Moldavian, as well as the reflexes given later for the same lexical items in Northern and Southern Moldavia, were taken from the following dialectal maps in the ALM (1968): *zi* [338]; *piept* [239]; *fačem* [474]; *meržem* [472]; *vulpe* [281]; *sjarè* [212]; *kyne* [77]; *luminè* [262]; *bèjat* [16]; *bèbat* [147]; *žos* [375]; *semn* [25]; *visez* [152]; *sè mjargè* [211]; *umblu* [504]; and *frate* [184]. All forms are given in phonetic transcription.

⁸With regard to this form, there is variation even within the group of dialects which encircle Kišinev. The reflex attested at site 140 [Kolonica] is *imbl*, a form which exhibits loss of final *-u*. This differs from *imblu*, the reflex found in the remainder of those dialects (i.e., sites 132, 133, 139, 145, and 148).

⁹Interestingly, such a claim depicts for literary Moldavian a scenario which is the exact opposite of one that applies to another Romance language, Raeto-Romance. Whereas Moldavian is essentially a literary language which finds nowhere among its local dialects a norm for speech, Raeto-Romance has standardized all five of its local dialects, each, today, being considered an accepted norm (Borodina 1969:11). I thank Eric Hamp (personal communication) for this observation.

¹⁰The semantics of *dînsul* may vary from dialect to dialect, and, in some cases, from speaker to speaker. As is the situation subsequently described for Moldavian, in Daco-Romanian, for certain speakers, *dînsul* may possess a degree of politeness only equal to that of *el*. What is clear, is that a pronominal form exists for both languages which possesses a third (or second, depending on the relationship between *el* and *dînsul*), highest degree of politeness. These are for Daco-Romanian and Moldavian the forms *dumnealui* 'he,' *dumneaei* 'she,' and *dumnealor* 'they' (Cazacu et al. 1980:245).

¹¹This lexical item, too, may be found in Schönkron (1952:31), but is entered as *buhaiu*. It is not characterized as "regional."

References

- ALM = *Atlasul lingvistik moldovenesk, Partja I și II*. 1968. Kișinev: Kartja moldovenjaskè.
- Borodina, M. A. 1969. *Sovremennij literaturnyj retoromanskij jazyk Švejcarii*. Leningrad: Nauka.
- Boršč, A. T., N. G. Korlètjanu, and E. M. Russev. 1954. *Russko-moldavskij slovar'*. Moscow: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej.
- _____, M. V. Podiko, and V. P. Solov'ev. 1961. *Moldavsko-russkij slovar'*. Moscow: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej.
- Bruchis, Michael. 1982. *One Step Back, Two Steps Forward: On the Language Policy of the Communist Party of the Soviet Union in the National Republics (Moldavian: A Look Back, a Survey, and Perspectives, 1924-1980)*. New York: Columbia University Press.
- Bugadov, R. A. 1953. "Moldavskij jazyk sredi romanskix jazykov (k postanovke voprosa)." *Voprosy moldavskogo jazykoznanija*, pp. 121-34.
- Cazacu, Boris, Matilda Caragiu Marioțeanu, Clara Georgeta Chiosa, and Valeria Guțu Romalo. 1980. *A Course in Contemporary Romanian*. Bucharest: Editura didactică și pedagogică.
- Čeban, I. D. 1953. "Sovremennoe sostojanie naučnoj razrabotki moldavskogo jazyka i ego istorii." *Voprosy moldavskogo jazykoznanija*, pp. 152-82.
- Diaconescu, Ion. 1977. *Infinitivul în limba română*. Bucharest: Editura științifică și enciclopedică.
- Dikcionar ortografik*. 1978. Kișinev: Redakcija prinčipalè a enčiklopediej sovetiče moldovenešt'.
- Dyer, Donald L. 1985. "The Interplay of Subjunctive and Infinitive Complements in Romanian." *Folia Slavica* 7:362-80.
- Filin, F. P. 1962. *Obrazovanie jazyka vostočnyx slavjan*. Leningrad: Izdatel'stvo Akademii Nauk SSSR.
- Gheție, Ion. 1964. "Cu privire la repartiția graiurilor dacoromâne. Criterii de stabilire a structurii dialectale a unei limbi." *Studii și cercetări lingvistice* 15:317-46.
- Gramatica limbii române, I, Fonetică și morfologie*. 1963. Bucharest: Editura Academiei Republicii Populare Romîne.

- Graur, Alexandru. 1955. *Studii de lingvistică generală*. Bucharest: Editura Academiei Republicii Populare Romîne.
- Il'jaŝenko, T. P. 1983. *Formirovanie romanskix literaturnyx jazykov. Moldavskij jazyk*. Kiŝinev: Ŗtiinca.
- Iordan, Iorgu. 1921. "Lexicul graiului din sudul Moldovei." *Archiva* 28:2:186-202.
- Korlètjanu, N. G. 1953. "Zadači issledovanija osnovnogo slovarnogo fonda i slovarnogo sostava moldavskogo jazyka." *Voprosy moldavskogo jazykoznanija*, pp. 183-206.
- _____. 1966. "Moldavskij jazyk." *Jazyki narodov SSSR, I. Indoevropskie jazyki*, 528-61, Moscow: Nauka.
- _____. 1969. *Limba moldovenjaskè literarà kontemporanè, Vol. I, Leksikoložija*. Kiŝinev: Lumina.
- _____. 1970. *Limba moldovenjaskè literarà kontemporanè, Vol. II, Fonetika ŝi morfoložija*. Kiŝinev: Lumina.
- _____. 1983. *Moldavskij jazyk segodnja*. Kiŝinev: Ŗtiinca.
- Limba ŝi literatura moldovenjaskè*. 1954-. Kiŝinev: Akademija de ŝtiincè a RSS Moldoveneŝt', Institutul de limbè ŝi literaturè.
- Lunt, Horace G. 1952. *A Grammar of the Macedonian Literary Language*. Skopje: Državno Knigoizdatelstvo.
- Macrea, D. 1958. *Dicționarul limbii romîne moderne*. Bucharest: Editura Academiei Republicii Populare Romîne.
- Marioțeanu, Matilda Caragiu, Ștefan Giosu, Liliana Ionescu-Ruxândoiu, and Romulus Todoran. 1977. *Dialectologie română*. Bucharest: Editura didactică ŝi pedagogică.
- Neiescu, P. 1971. "Cu privire la repartiția dialectală a dacoromânei." *Actele celui de al XII-lea Congres internațional de lingvistică ŝi filologie romanică, vol. I-II*, pp. 297-302.
- Philippide, Alexandru. 1925-28. *Originea romînilor, t. I-II*. Iași: Viața romînească.
- Rusu, Valeriu. 1968. "În legătură cu structura dialectală a dacoromânei (Pe marginea unor studii recente)." *Studii ŝi cercetări lingvistice* 19:5:521-23.
- _____. 1984. *Tratatul de dialectologie românească*. Craiova: Scrisul românesc.
- Schönkron, Marcel. 1952. *Rumanian-English and English-Rumanian Dictionary*. New York: Frederick Ungar.
- Sergievskij, M. V. 1959. *Moldavo-slavjanskije ètjudy*. Moscow: Akademija Nauk SSSR.
- Šiŝmarev, V. F. 1953. "Romanskije jazyki jugo-vostočnoj Evropy i nacional'nyj jazyk Moldavskoj SSR." *Voprosy moldavskogo jazykoznanija*, pp. 73-120.
- Talmackaja, M. F. 1978. *Moldavskoe jazykoznanie v sovetskoj pečati*. Kiŝinev: Ŗtiinca.
- Tiktin, H. 1903. *Rumänisch-deutsches Wörterbuch (Dicționar român-german), t. I*. Bucharest: Staatsdruckerei.
- Vasiliu, Emanuel. 1968. *Fonologia istorică a dialectelor dacoromâne*. Bucharest: Editura Academiei Republicii Socialiste România.
- Vulpe, Magdalena. 1963. "Repartiția geografică a construcțiilor cu infinitivul ŝi cu conjunctivul în limba română." *Fonetică ŝi dialectologie* 5:123-55.

- Weigand, Gustav. 1909. *Linguistischer Atlas des dacorumänischen Sprachgebietes*.
Leipzig: Johann Ambrosius Barth.

Complex Infinitives and Other Deverbal Nominals in Tajik

Rachel Lehr

There are two types of complex structures in Tajik which function as deverbal nominals. Traditional grammars take these forms to be essentially equivalent, stating that the difference is that one is a phrase and the other a word. The transparency and productivity of complex morphological expressions in Tajik has been recognized for its systematicity and predictability all of which bears on the distinction of word and phrase. Thus in looking at these two complex strategies for forming deverbal nominals we will examine their source — complex verbal expression — and argue that their difference, if any, stems from this source. We will also compare and evaluate the semantic, morphological, and syntactic differences to determine the function these two types of nominals serve in the language as a whole.

1. In our first example we have nearly identical sentences;¹

- 1a. dars xondan-i hasan befoida-st.
lesson read (inf.)-ezf. Hasan useless-3sg.cop.
- 1b. dars-xoni-i hasan befoida-st.
lesson read-ezf. Hasan useless-3sg.cop.
'Hasan's studying is useless.'

In example (1a) we have a verb in the form of the infinitive functioning as an action nominal. In example (1b) the function of the verbal element is also that of an action nominal but its form is bound and does not occur independently. We will return to this formal difference in the next section; for now we would like to look at the semantic differences between (1a) and (1b).

The semantic difference between the two forms represented by the sentences in example (1) are immediately apparent to any speaker of the language, although this difference is not always distinctive. Examples of the type found in (1b) (darsxoni) refer to actions which are generally performed in a habitual or professional manner and often may be used to denote the product of such actions. The example in (1a) (dars xondan) and others like it may be used to refer to a particular instance of this

activity, either occasional or regular. The infinitive does not carry the assumption of repeated or professional action, although it may at times be used that way. In contrast, the example in (1b) has only that meaning and may not be used to refer to an occasional instance. The following examples illustrate this contrast:

- 2a. pas az tašnagi xalos {kardan-i } asp-ho-yašon boz ba
 after thirst free do { (inf)-ezf. } horse-pl-3pl.poss. again to
 { *kuni-i }
 roh daromad-and
 road enter 3 pl.
 'After the thirst of their horses had been quenched, they started out on the road.'
- 2b. { šarob-xuri } qin-i kalon ast
 wine consume { problem-ezf. large 3sg.cop.
 { ??šarob xurdan }
 wine consume (inf.)
 'Alcoholism is a big problem.'

In (2a), from Rastorgueva (1963:98), the action is clearly discrete, neither habitual nor professional and the infinitive is preferred. In (2b) the action is a general condition requiring repeated consumption to qualify and the infinitive is far less acceptable.

It might be said that there is a perfectiveness associated with the noninfinitive form such that the action being repeated either habitually or professionally is assumed to be completed.² In the next example we see that while both forms may be used to refer to products of action (although this is more limited for infinitives), the infinitive implies work in progress while the other form is for a completed product.

3. { gul duxton-i } in zanak xebe zebo-st
 flower sew { (inf)-ezf. this woman very pretty-sg.cop.
 { gul-duzi-i }
 flower-sew-ezf.)
 'This lady's embroidery is very pretty.'

The sentences in (3) may refer either to the act of embroidery or to the product itself. Adding the plural suffix *-ho* to the forms in (3) has the effect of concretizing and the nominals refer to products:

- 4a. gul duxtan-ho-yi in zanak...
 flower sew (inf.)-pl-ezf.
 4b. gulduzi-ho-yi in zanak...
 flower sew-pl-ezf.

There is a difference however in that in (4a) there is a sense of unfinished work, in (4b) the products are considered complete.

To summarize our observations so far, we have seen that both of these complex nominal forms have a range of meanings but are also semantically in complementary distribution. The infinitive is used for any action, unique or repeated although not generically, the other form is reserved for routinized action. Both may be concretized as products but the infinitive for incomplete products, the other for finished products. These differences are displayed in the following table:

	generic action	discrete action	product
infinitive	-	+	-complete
noninfinitive	+	-	+complete

2. In this section we will examine the structure of complex verbs in Tajik. This is the obvious place to turn in examining complex forms that are deverbal in nature. In Tajik there are a limited number of simple verbs commonly used. Increased expression of verbal need is through use of more complex verbal expressions. These expressions generally consist of a verb and preverb, usually an adverb, adjective, preposition, or noun. The types of such verbal forms have been characterized according to the syntactic, semantic, and morphological cohesion of the verb and preverb. These complex expressions are generally discussed under the rubric of *compound verb*.³ Compound verbs in Tajik are usually characterized by the relation between the two elements. If the preverb is a noun, it is obligatorily generic, lacking marking for number as well as the postposition for specific direct objects. Between preverb and verb nothing appears beside verbal morphology and a prenominal enclitic.⁴ The question naturally arises, however, whether all structures meeting these criteria are compound verbs. For example:

5. Hasan kitob me-furuš-ad
 Hasan book cont.-sell-3sg.
 'Hasan sells books.'

This sentence fits the structural description of compound verbs and is a generic activity. However, the object may easily be specified with a demonstrative and a postposition as in (6) below. It is clear that the sentence is different but the meaning of the verb is basically unchanged.

6. Hasan in kitob-ro me-furuš-ad
 Hasan this book-d.o. cont.-sell (pres.) 3sg.
 'Hasan is selling this book.'

There is a tendency for generic objects to join with the verb for routinely occurring, general activities. Material intervening between the verb and its object is generally excluded. However, there are some additional characteristics by which compound verbs are identified. Like incorporated nouns found in many other languages, the preverb lacks the syntactic status of an argument in compound verbs in Tajik. As such, preverbs serve the function of modifying the action described by the verb. The occurrence of many adverbs and adjectives in compound verbs supports this.⁵

Adverbs and adjectives have relative positions in the sentence. Adverbs of manner are usually put before the verb, though not exclusively. In the case of a compound verb, the adverb is fixed in position with the verb. Consider the following example.

- 7a. Hasan in kitob-ro ohista xond
Hasan this book-d.o. slowly read (past) 3sg.
'Hasan read this book slowly.'
- 7b. in doxtar-ro naghz me-bin-am
this girl-d.o. well cont.-see (pres.) +sg.
'I like this girl.'

In both sentences in this example the adverb occurs just before the verb. Only in (7a) may the adverb move its position. Movement of the adverb in (7b) results in an ungrammatical sentence. This is illustrated below in example 8.

- 8a. Hasan ohista in kitob-ro xond.
- 8b. *naghz in doxtar-ro me-bin-am.

We find the following observations in Rastorgueva (1963:95) concerning the order of constituents in a Tajik sentence: the direct object without postposition *-ro* usually stands immediately before the verb. When the postposition *-ro* is present the position of the direct object is freer and can be separated from the verb by an indirect object or by an adverbial of place or manner. The following examples from Rastorgueva (1963:95) illustrate this difference.

- 9a. man az daryo ob ovar-d-am
I from river water bring (past)-1sg.
'I brought water from the river.'
- 9b. in kitob-ro az magazin xarid-am
this book-d.o. from shop buy (part)-1sg.
'I bought this book in a shop.'

Let us turn now to the nominal preverb in compound verbs. In the overwhelming majority of compounds, the direct object is joined with the verb. In addition to direct object we find locatives and instruments, some with overt prepositions and some without. Looking closely at this data we will argue that none of the preverbs maintains an argument position in the sentence.⁶ What is or is not a syntactic argument is a widely contested question which we will not address in this present discussion. The most obvious candidate for an argument of a verb is direct object and this is what interests us in Tajik. Traditional grammar considers the nominal preverb in the majority of compound verbs to be direct object, presumably because of semantic interpretation. Notably, there is no preposition and, as mentioned earlier, the postposition *-ro* is absent. The absence of *-ro* is traditionally seen as due to the generic nature of the preverb, *-ro* being used only for specified objects. I would like to offer an alternative explanation. The absence of *-ro* is due to the fact that this preverb is not functioning as a syntactic direct object at all. We present here evidence of two sorts. First, we have sentences with an overtly marked direct object outside the compound. For example:

10. vay-ro dust me-dor-am
 s/he-d.o. friend cont. -have (pres.) 1sg.
 'I like him/her.'

Second, we will look at a test used by Moyne (1970:51) and further discussed in Windfuhr (1979:120) as a method for identifying compound verbs in Persian. In this test the nominal preverb is juxtaposed with the verb and put into an *ezafe* construction. Verbal expressions which pass this test are not compounds since compound verbs are of fixed structure and are generally lexicalized at least to the extent that juxtaposition would change the meaning or yield ungrammatical results.

	infinitive	ezafe construction
11a.	dars xondan lesson read 'to study'	xondan-i dars reading-ezf lesson 'to read a lesson'
11b.	roh raftan road to go 'to walk'	*raftan-i roh
11c.	kitob furuxtan book to sell 'to sell books'	furuxtan-i kitob selling-ezf books 'to sell books'

Palmer (1971) notes that a transitive infinitive in Persian may be in *ezafe* construction only with its direct object and not with its subject or any other syntactic argument. The following example in Persian shows this:

12. **xordan-e Hasan*
 consume (inf.)-*ezf. Hasan*
 'The drinking of Hasan.'

In (12) the transitive verb is construed with its subject and the result is ungrammatical. Likewise, in (11a) the 'direct object' preverb does not yield the same meaning in either construction. In (11b) an intransitive verb where there can be no direct object is ungrammatical in *ezafe* construction. All of this is in contrast to (11c) where both constructions have the same meaning.

In this section we have looked at some of the special features which characterize compound verbs. We have been able to differentiate true compound verbs from those constructions with similar structures. We have shown that a unique relation obtains between nominal preverb and the verb in compounds in which the lack of syntactic argument status results. The infinitivization test of Moyne (1970:51) was discussed in connection with this. This test was used by Moyne to identify whether the infinitive is part of a compound verb or not. Examples (11a) and (11b) as opposed to (11c) show the different results. The elements of compound verbs cannot be juxtaposed into *ezafe* construction without a change of meaning or becoming ungrammatical. The category of compound infinitive, not previously discussed in this paper, will be a focal concern in the following section of this analysis.

3. Now we return to the question posed in the first section of this paper: What is the difference between the two types of complex deverbal nominals examined in section 1? We have already noted the semantic difference between the constructions. Let us turn to the morphological difference.

The infinitive is an independently occurring word and its form does not change whether in compound or not. The other nominal, however, is different. It is always a bound form and systematically occurs non-independently. Thus, the only word definable here is either the preverb alone or the compound as a whole.

Orthographically there is a tendency to write the compounds as a single word and the compound infinitive as separate words. Stress assignment does not tell us much. Stress is generally word final and compounds are pronounced with main stress on the final syllable, and so,

in fact, are phrases. According to Rastorgueva (1963:100) both compounds and compound infinitives are pronounced as a single utterance.

If both compounds and compound infinitives originate in the compound verbs discussed in the second section, why should there be a difference in morphological or syntactic status? We have already seen that there is a morphological difference such that one form is bound and the other is not. Let us look now at how the syntax treats them.

Both forms are clearly nominal. We observed earlier (example (4)) that they both take the plural suffix *-ho*. They both can be modified by demonstratives:

13. $\text{hamin} \left\{ \begin{array}{l} \text{sigar kašidan} \\ \text{sigar-kaši} \end{array} \right\} \text{-at ba jonam rasid.}$
 this cigarette smoke-2sg. to soul-1sg reach
 ‘I don’t like your cigarette smoking.’

Modification of items internal to the compounds is the same as to the compound infinitive. The first element may be modified as long as the meaning does not become referential or specified. In the above example we could not interpret *hamin sigar-kaši* as ‘your smoking of this cigarette,’ but rather we interpret it as ‘this cigarette smoking of yours.’ Insofar as the infinitive is interpreted as a compound, it will have the same meaning.

In example (14) we have the first element modified internally:

14. *kitob-i-tojiki-furuši*
 book-ezf-Tajik-sell
 kitob-i tojiki furuxtan
 book-ezf Tajik sell (inf.)
 ‘selling Tajik books’

Is there then a difference of word status between these two forms? Because of the morphological character of the compound it is clearly definable as a single word. The compound infinitive is more elusive. We have seen that the cohesiveness between the preverb and the verb is related to a number of factors and although the number of compound verbs is great they are predetermined and lexicalized in meaning. There is, however, a tendency for verbs to coalesce syntactically and semantically with generic preverbs to denote generic, nameworthy activities and this process can be seen as innovative. The problem we have in determining the status and function of infinitives in their nominal use is in knowing to what extent the infinitive is compounded or stands alone. The infinitive is clearly in compound form when it corresponds to a predetermined compound verb and then its parts are as inseparable as

those of the compound. In this case the difference between the two forms would be semantic. Since we also have complex infinitives which do not correspond to compound verbs (although perhaps tending in that direction) we expect to and do find a syntactic difference.

The order of arguments incorporated in compounds is fixed. Transitive verbs incorporate direct objects or modifiers first, not indirect objects. For example:

- 15a. dar ošxona sigar-kaši / sigar kašidan
in kitchen cigarette-smoke/cigarette smoke (inf.)
- 15b. ba kudak-on tuhfa-dihi / tuhfa dodan
to child-pl. gift give / gift give (inf.)
'gift giving to children'

In both of the examples in (15) the direct object occurs closest to the verb. Reversing the order is not possible for the compound; neither is incorporating the indirect object and putting the direct object outside:

- 16a. *dar ošxona kaši-i sigar
- 16b. *ba kudakon dihi-i tuhfa

The infinitive when not compounded is not restricted in this way:

- 17a. dar ošxona kašidan-i sigar.
- 17b. ba kudakon dodan-i tuhfa.

Conclusions. We have seen that compound infinitives can be identified insofar as they correspond to compound verbs. The same is not true of the other compounds. Compound nominals such as *kitobfuruši* 'bookselling', *darxoni* 'studying', *gulduzi* 'embroidering', and others belong to a fully productive class. The morphological process creating these forms is limited only by the speaker's imagination. More important though is that the lexical items produced in this way need not correspond directly to known compound verbs. Many innovative forms are found in speech and particularly in journalistic writing. This is one of the strategies in Tajik used to create a new name for any generic activity. In contrast, the infinitive, whether simple, compound or complex is not used in this naming sense but is rather an action nominal of general use. We would suggest then that whether the infinitive is in a compound construction or not is not an important part of its word status or function. When the infinitive is part of a compound it is not different morphologically from the other compound form but its source perhaps is of a lower level morphological process, that of compound verbs. The word status of the compound nominal is never in doubt because of its

bound form. We would like to suggest that its source is not in compound verbs but in a higher level, more productive morphological process.

University of Chicago

Notes

¹In glosses to examples, the following abbreviations are used: *cop.* copula, *cont.* continuous aspect, *d.o.* direct object, *ezf.* ezafe, *inf.* infinitive, *pl.* plural, *poss.* possessive, *pres.* present tense, *sg.* singular.

²Carlson and Roeper (1980:136-37) speculate that complex verbs, in particular prefixed ones, might have the semantic property of perfectiveness. While they do not pursue this possibility very far in their account it does seem to be a common property of deverbal compound nominals in Tajik.

³The compound verb is a much discussed area of Persian and Tajik linguistic investigation. Opinions range from the nonexistence of this category to its subdivision into several categories. No two researchers agree.

⁴The view presented here follows Moyne (1970) and many standard grammars. It is by no means conclusive. Bashiri (1973:148-51) shows a number of places where other material can intervene and Telegdi (1951:326-30) argues that intervening material does not change the status of a compound. At the root of this controversy is the notion 'compound' itself.

⁵In many cases adverbs and adjectives are not formally distinguished in Tajik and so we include both as preverbs. It is clear that the function is adverbial.

⁶Mithun (1984) discusses how incorporated nouns typically lose their syntactic argument status. In type I noun incorporation the valency of the verb is reduced. In type II another argument may occupy the vacated case role. Both these types seem to be present in Tajik. In classical Perso-Tajik the case marker *-ro* was used for both direct and indirect objects. In the contemporary language its use with the indirect object is obsolete. With compound verbs, where the direct object is incorporated, an oblique argument is marked with *-ro*; for example:

zabon-i tojiki-ro dars me-dih-am
language-ezf. Tajik-d.o. lesson cont.-give-1sg.
'I teach Tajik.'

References

- Bashiri, I. 1973. *'To Be' as the Origin of Syntax*. Minneapolis: Bibliotheca Islamica.
 Carlson, G. and T. Roeper. 1980. "Morphology and Subcategorization: Case and the unmarked complex verb" in *Lexical Grammar*, ed. T. Hoekstra et al., Dordrecht: Forist, 123-64.
 Mithun, M. 1984. "The Evolution of Noun Incorporation." *Language* 60:847-94.

- Moyne, J. A. 1970. The Structure of Verbal Constructions in Persian. Harvard University dissertation.
- Palmer, A. 1971. The Ezafe Construction in Modern Standard Persian. University of Michigan dissertation.
- Rastorgueva, V. A. 1963. *A short Sketch of Tajik Grammar*. Trans. and ed. by H. Paper. The Hague: Mouton.
- Telegdi, Z. 1951. "Nature et fonction des périphrases verbales dites 'verbes composés en Persan'." *Acta Orientalia*. 1:315-38.
- Windfuhr, G. 1979. *Persian Grammar: History and State of its Study*. The Hague: Mouton.

Stressing Freely in Lithuanian and Russian

Jules F. Levin

...My own experiments taught me that the influence of Russian musicality is always harmful: when the ear is sensitive to a language with strong accents, the natural tendency to rival it can be disastrous for weakly accented languages like Polish or Czech. The precision of line and of intellectual nuance that can be achieved in these languages is then destroyed by the wild impulse to write a tribal song or by a measured beat — in that case the metrical ideal of all poetry would have to be E. A. Poe's *The Raven*, and it surely is not.

— Czesław Miłosz, *Native Realm: a Search for Self-Definition*. 1968.
(Garden City, NY), p. 123.

It is more important than ever to clarify the perspective of any linguistic analysis. In addition to the traditional dichotomies of diachrony and synchrony, the latter itself can be considered in several ways. In recent years most synchronic descriptions have been speaker-storage and production models. The algorithms that mediate between “representations” and surfaces have as their targets the output of **speakers**. Synchronic **acquisition** models, though still considered more psycholinguistics than pure linguistics, are not necessarily identical with speaker-storage and production models. There is yet another possible synchronic structural perspective, that of the listener. Listening and comprehending speech do not necessarily involve mere storage and production models played backwards; the tasks involved demand specific cognitive operations that are not present in production. Language evolves through the interaction of speakers and hearers in countless speech acts, and can be viewed as a dialectic, since the interests of speakers and hearers do not always coincide. The problem of recovering

a message from the complex phonic information entering the brain is so great that the brain uses every possible clue available to it.

In the perceptual model phonetic detail comes into its own: Aspiration, partial devoicing, vowel reduction, etc., provide valuable clues to otherwise elusive morpheme and word boundaries. Phonological structure limitations, such as those first noted by Jakobson (1984:119) for the Russian desinential system, help the brain exclude possibilities: [róskəʃ] can only be a stem, whereas [róskəm] or [róskəf] may consist of stem + desinence.

Neutralization, which in storage and production models often seems to have a purely formal, if not arbitrary character, suddenly emerges as overtly functional. Consider classical Praguian neutralization, like the loss of voicing contrast in final position. To appreciate the function of such neutralization for the speech perception model, one must recognize that a **position of neutralization** is an unmarked position, and therefore ambiguous. What is not neutralized *marks* the position where neutralization does not occur. In Russian, a voiced consonant followed by a vowel or sonorant is *not* in word-final position, therefore the listener knows that he has not yet heard the end of the word. This analysis yields an empirical definition of markedness: A marked sign is a sign that must exclude a possible interpretant. This is a listener-oriented definition.

* * *

In typological discussion of stress, Lithuanian (Lith) and Russian (Russ) are usually grouped together among languages with free stress. In such languages stress can fall on any syllable, and is phonemic — potentially functioning as a differentiating sign at the phonological level. However, comparisons usually go beyond a typological similarity. Stress in Lith and stress in Russ reflect the diachronic development of stress in Baltic and Slavic respectively, and to the extent that stress in the reconstructed ancestors is genetically related, so that Slavic stress can be shown to have a regular neogrammarian-style relationship to stress in Baltic, stress in Lith and Russ are also genetically as well as typologically taken as variations on a single theme; two ways of being the “same thing” rather than two different things. It is not surprising therefore, that analyses of stress in Lith that seek some deep-level understanding, also strive to find in that stress a deep-level unity with the deep-level nature of stress in Russ. For example Kenstowicz in his dissertation on Lithuanian phonology points out parallels with Slavic: “This permits the generalization that all masculine singular desinences are stressless, which of course corresponds to the situation in Slavic,” then concludes that: “...it would

appear that the basic underlying pattern in the distribution of stress in the desinences is one of an opposition between the masculine ... and feminine ... in the singular and a neutralization of this opposition in the plural...." (Kenstowicz 1972:77). This conclusion facilitates comparison with Slavic.

Superficially this attempt seems perfectly reasonable. Consider the following well-known comparisons.

	Nom sg	Gen sg	Acc sg	Nom pl	Gen sg
R	golová	golový	gólovu	gólovy	golóv
L	galvà	galvõs	gálva	gálvos	galvũ
	1st P sg		2nd P sg		3rd P sg
R	vížu		vídiš		vídit
L	dúodu		dúodi		dúoda
R	mogú		móžeš		móžet
L	kerpù		kerpì		keřpa

Given such parallels why would anyone try to find underlying divergence rather than underlying unity? We make the attempt out of perversity, and our approach is from the listener's perspective.

Stress seems to be a listener's function par excellence. Perhaps that explains why stress information is so often omitted in writing systems. Stress has some function in every language. Even in French, stress marks the end of a logical phrase unit. In languages with fixed word stress, such as Latvian, stress serves to mark word boundaries. A stressed syllable in Latvian means "new word starts here."

On the other hand, stress in languages with free stress can have differing functions. One can divide stress function into two types — syntactic and derivational (paradigmatic). Derivational function operates at the level of lexical and semantic organization in the brain. Let us imagine a language in which all simple root nouns and verbs had only fixed end stress and all derivatives had only fixed stem stress (i.e., Lith *rankà* ~ *rankėlė*; Russ *ruká* ~ *rúčka*). In such a language a stressed desinence would simplify the search procedure of the listener in matching the perceived form and an appropriate meaning. The listener can automatically exclude from the search derived lexicon. Other kinds of derivational function similarly assist in the semantic identification of words in the speech chain — for example tendencies to distribute stress assignment by gender. All stress limitations that separate classes of words serve to speed up listener interpretation of the message through simplifying lexical searching.

On the other hand, other kinds of stress distribution help the listener interpret the syntax of a message. For example, consider a language with initial stress on all verbs and final stress on all nouns. Any non-initial stress perceived by the listener would exclude the interpretant "verb." In the flow of speech this constitutes syntactic rather than derivational information.

It is clear where Russian falls in this division. In Russian stress can fall on the root or on the ending for any noun or any conjugated verb form, and therefore stress can only serve to distinguish lexical meaning. This is the only possible function of stress in Russian. That is, one finds the minimal stress contrasts along the paradigmatic axes, not along the syntagmatic axes. Avanesov states that

"Positional variation of stress makes stress in Russian an *individual attribute* of every separate word. Every word has its appropriate stress on this or that syllable. One on the first syllable, another on the second.... In this regard the place of stress, being an attribute of the given word, relates to the area of the lexicon — to the inventory of means serving to differentiate words" (Avanesov 1958:18).

Even if we turn to stress alternation within the Russian paradigm, there also contrastive function is limited to the particular word nest: Nom pl *gólový* with stem stress can contrast only with Gen sg *golový* as an internal paradigm contrast (like the alternation /z ~ r/ in English *was* ~ *were*, which does not regularly signal number contrast in the verb.) In Russian there are plenty of Nom pls in stressed -ý, and of course plenty of Gen sgs with a stem-stressed -y ending.

But what about Lithuanian? In contrast to Russian, Lith is characterized by a limitation of stressing possibilities on the syntactic plane. Paradoxically, it is the presence of such a limitation that permits stress to have a communicative function in syntax. Indeed, because stress is so often looked at in the context of the traditional vertical paradigm, the true function of stress in Lith is obscured, and one can only see a false parallelism with Russ.

The most important stress limitations in Lith are the following:

- (1) In the verb paradigm, the crucial fact of stress is that a contrast of stress between root and ending is possible only in the first and second person singular. Elsewhere only stem stress is possible; i.e., third person and pl are positions of neutralization for stress.
- (2) In Nominal declension no Accusative or Dative sg can have desinential stress; i.e., stress is neutralized in the Acc and Dat sg.

(3) There is a tendency for stress to be fixed on the stem in the Nom case.

We can now return to the themes of our preliminary remarks. Stress alternations between stem and ending are excluded from particular grammatical functions — the third person (and plural verbs), the Acc and Dat sg. These are positions of neutralization, and therefore *marked stress* — ending stress — excludes these functions as possible interpretants, a valuable clue for the brain trying to make sense of the flow of speech.

Consider the implications of the Acc case stress rule — no Acc sg ending can be stressed. This limitation must be a valuable aid in sentence comprehension — no word with ending stress can be interpreted as the sg direct object of a verb. (And parenthetically, a stressed stem followed by a long vowel is very likely to be an Acc sg.)

Contrast this with Russ, which has stress retraction in the Acc sg for some *-a* stem nouns (in some cases cognate with Lith nouns). However, since an Acc sg ending may or may not be stressed, this stress retraction serves no function in interpretation on the syntactic plane. A stressed *-ú* can be a possible fem Acc ending.

What is the structural significance of the stress limitations which we have in Lith? It seems that taking all the facts of stress limitation together, what is most important is that Lith either neutralizes or tends to neutralize the contrast between stem and ending stress in what we call **unmarked syntactic structure**. Consider a sentence like:

- (1) Jōnas dāvē meŕgai šūnį.
'Jonas gave the girl a dog.'

Such a sentence superficially "feels like" a similar Russian sentence:

- (2) Iván posyláet dévuške sobáku.
'John sends the girl a dog.'

But in the Lith sentence there is no "free" stress: all stresses are determined by stress neutralizations! Note that the Dat sg also is included in this central unmarked syntax. On the other hand, compare a sentence like

- (3) Aš rašaũ sàvo rankà laišką kambarĩ.
'I wrote a letter in the room with my own hand.'

End stresses characterize "marked" categories — 1 sg., Inst, Loc, and therefore exclude "unmarked" categories as interpretants. A similar Russ sentence:

(4) *Vesnój ja pišú pís'ma v sadú.*

'In the spring I write letters in the garden.'

does not exclude any case forms as possible interpretants, compare:

(5) *Oná pošlēt pís'mó sestré.*

'She will send a letter to her sister.'

A similar sentence is not possible in Lith.

The syntactic function of stress for the listener can be determined by observing where contrast is possible, and where it is not. Its main function in this context must be to exclude interpretative possibilities from the syntactic plane. In Lith we find the possibility of stress contrast between stem and ending in marked, peripheral syntax. When stress falls on the ending in these forms its function is to exclude one interpretative possibility, to simplify the syntactic interpretative burden of the listener.

Because stress (but not tone/accent contrast) is neutralized in central unmarked syntactic structure, it seems unlikely that stress in Lith has a very important derivational/lexical (paradigmatic) function. Note that stress paradigms as such never contrast, never correlate as minimal pairs, **except in marked syntax**. In fact, the preservation of accent contrasts as a derivational/lexical function in Lith permitted the neutralizations of stress that result in the latter's syntactic function.

Tendencies in Lith seem to be in the direction of manifesting this syntactic function. Stress is moving to the stem in the Nom case, and to the endings in oblique cases. Steve Young (personal communication) reports such anomalies as Nom *Rýga*, Loc *Rygōj* 'Riga'. The markedness of stress position (in general, we assume ending stress to be marked vis-a-vis stem stress) iconically maps the markedness of category memberships.

It occurred to me to wonder if these differences between Russian and Lithuanian stress yield differences in the flow of speech. I have read through some examples of poetic speech and prose to see if the differences can possibly make a difference. I took as one comparison text the Lith and Russ translations of a passage from Hamlet (Act III, Sc. 4, lines 164-177). The metrics require that some lines end on a stressed syllable... One problem of comparison is what to do with stressed final stem syllables in Russian. We can take a form like *ljubón'* as having stem stress (cf. Inst *ljubón'ju*), and one like *pogloščěn* as having end stress, since the latter would have end stress if there were a non- \emptyset ending.

Argi tai méilė? Ne, ne tuo keliu
 Jo svajos vėžiasi, ir jo kalbà,
 Nors trūputį išdrikusi atródė,
 Nebuvo svaičiojimas. Jōjo sėloj
 Nusiminimas lizdą susisūko,
 Ir aš bijaū, kad jis neišperėtų
 Kokiōs bėdōs; todėl, kol ne vėlū,
 Aš nesvyrūdamos nusprėndžiau šitaip.
 Jis Ąnglijon nedėsdamos tevỹksta
 Išreikalaut nesumokėtos dúoklės;
 Gal kaftais jūra ir nauji kraštai
 Ir reginių margūmas išblaškỹs
 Mintis, kuriōs, užgūlusios jam širdį,
 Kamuoja smėgenis be atvangōs,
 Iškreipdamos jo tikrąjį pavėikslą.
 Kaip jūs galvójate?

Ljubov'? On pogloščēn sovsēm ne jeju.
 K tomú ž xot' svjázi net v egó slovax,
 V nix net bezúmja. On ne to lelėet
 Po tēmnyu ugołkám svojęj toskí,
 Vysízivaja čto-to poopásnej.
 Čtob vó-vremja bedú predotvratít',
 Prišel ja k slėdujuščemu rešén'ju:
 On v Ąngliju ne mēdlja otplyvėt
 Dlja sbóra nedovýplačēnoju dání.
 Byt' móžet, móre, nóvyje krajá
 I ljúdi výb'jut u negó iz sėrdca
 To, čto sidít tam i nad čém on sam
 Lomáet gólovu do otupén'ja.
 Čto dúmaete vy na etot ščēt?

In reading the two translations of Shakespeare over and over I began to develop a sense of differing “flavors” of stress. The first thing I noted was that ending stress seemed to have a stronger aesthetic potential in Russian. The fourth and fifth lines illustrate what I consider a very typical and powerful use of end stress in Russ that seems to be lacking in Lith. Although the number of end stresses is about the same, Lith seems proportionally to have fewer, and as I became conscious of the neutralization in Lith, many lines became less interesting with respect to stress, cf. lines 9, 10. Curiously the effect was to make Lith sound more like the English original, while the more powerful rhythmic use of end stress in the Russ distances it from the English metrical “flavor.”

After a while Lith sentences with neutralized stress, hence “unmarked” syntax, began to sound almost like Polish. Shifting the main stress function from lexical to syntactic through partial neutralization may be a path towards fixing stress on a particular syllable, i.e., to reducing stress functions to marking word boundaries.

I also compared an Axmatova poem with its translation into Lith by Venclova. At first it was hard to notice any difference in stress use. Both versions had plenty of stressed endings, but since the Russian did not use stressed Acc or 3rd person forms the stressed endings were virtually the same. But two lines particularly struck me. The Russian reads (Axmatova 1977:61)

Ty dýšiš' sólncem, ja dyšú lunóju,
 No žívy my ljubóviju odnóju.

In the first line stem stress in the first half contrasts with end stresses in the second — i.e., *dýšiš' sólncem* contrasts with *dýšú lunóju*, and this reinforces — iconically maps — all the other contrasts — 1st person ~ 2nd person, female ~ non-female, sun ~ moon. The next line with root stress on *žívy* and ending stress on *odnóju* echoes the contrast. Venclova (1979:12) translates these two lines:

Tau švýti sáulė, ménuo man bolúoja,
Bet víena méile síelos tealsúoja.

Everything is stem-stressed! Whatever else is going on, for Venclova the aesthetically pleasing semiotic use of stress is not a viable option here. Does the syntactic function of stress in Lith restrict its use as a poetic device compared to Russian?

University of California, Riverside

References

- Axmatova, Anna. 1977. *Stixotvoreníe i poëmy*. Leningrad.
- Avanesov, R. I. 1958. *Udareníe v sovremennom russkom literaturnom jazyke*, 2nd ed. Moskva.
- Jakobson, Roman. 1984. *Russian and Slavic grammar: studies 1931-1981*. Berlin, New York, Amsterdam.
- Kenstowicz, Michael. 1972. "Lithuanian phonology." *Studies in the linguistic sciences* (Department of Linguistics, University of Illinois, Urbana-Champaign) 2,2:1-85.
- Shakespeare, William. 1941. *Gamlet, princ datskij*. (Trans. Boris Pasternak.) Moskva.
- _____. 1970. *Gamletas*. (Trans. Aleksys Churginas.) Vilnius.
- Venclova, T. 1979. *Balsai*. Southfield, Michigan.

The Original Geographic Distribution of the Tungus Languages

Roy Andrew Miller

In Memoriam: V. I. Cincius (1903-1983)

The Tungus languages not only constitute one of the most substantial, and widely diversified, non-Slavic linguistic stocks existing in the USSR today, but they are also of considerable importance for the general comparativist, partly because of the internal evidence that they preserve for the earlier existence of a proto-Tungus (pTg) linguistic unity, but even more because of the essential role that these languages must inevitably play in any and all attempts toward the reconstruction of proto-Altaic (pA), that still more remote *Spracheinheit* whose existence we must postulate if we are to account historically for the innumerable and striking parallels in matters of phonological, morphological, and lexical detail that comparison between Turkic and Mongolian on the one hand, and Tungus on the other, reveals. The critical role of Tungus data in the reconstruction of pA has long been recognized by Altaicists, and attempts at the recovery of critical pA forms and formations largely documented by the evidence of Tungus cognates have long played an important role in the Altaic literature, even (it must now be admitted) in a few notorious instances where the critical Tungus forms were in effect linguistic ghosts, the result of misquoting of the at best fragmentary accounts of these languages which until recently were all that was available to our science (e.g., the Nanai ghost-word **balža* 'head', on which see now Miller 1984:159).

Today, however, and thanks almost entirely to several decades of most impressive scholarly labor on the part of our Soviet colleagues, it is no longer necessary, when discussing Tungus, and most importantly of all, when introducing Tungus linguistic data into the problems of Altaic comparison and reconstruction, to rely exclusively upon such antique, if in considerable measure still important, sources as Grube 1900. The entire face of Tungus comparative linguistics was altered by the publication of Cincius 1949, and it was the rich body of comparative materials so painstakingly assembled by this late Soviet scholar that

formed the principal source for the useful and frequently cited handbook-like monograph of Benzing 1955. But it remained for the publication between 1975 and 1977 of the two volumes of the long-awaited *Sravnitel'nyj slovar' tunguso-man'čžurskix jazykov* (=TMS), edited by Cincius with a number of her Leningrad colleagues, to make easily available to us all the enormous fund of precious lexical materials so carefully and diligently assembled by the Leningrad school over decades of fieldwork, frequently conducted under conditions of considerable difficulty and often, too, personal hardship. With the two volumes of TMS in hand, we are now at last in a position to initiate serious and responsible comparative studies of this important non-Slavic linguistic stock of the Soviet Far East.

In TMS we have at our disposal reliable linguistic data for ten different Tungus languages. These comprise (a) four modern literary languages: (1) Evenki (*ëvenkijskij*), Ev., the language frequently cited, especially in the earlier literature, simply as 'Tungus,' in the sense that it is, or may be considered, 'Tungus proper,' *eigentlich tungusische*; (2) Lamut (*ëvenskij*), La.; (3) Nanai, Na., in the earlier literature generally 'Goldi'; and (4) Udihe, Ud.; (b) one now extinct literary language, (5) Manchu, Ma.; (c) five contemporary spoken languages without standard written forms, (6) Solon, Sol.; (7) Negidal, Neg.; (8) Oroči, Oroč.; (9) Oroki, Orok., and (10) Olča or Ulča (*ul'čskij*), Olč.

In Tungus, earlier written records are available only for Manchu, both in its own right and also in the form of a small but important corpus of slightly older forms attested in Chinese transcription for an earlier variety of Manchu generally cited as Jurchen, or Žürčen (Jr.). The TMS most usefully includes Jurchen cognates whenever its editors have been able to identify them; otherwise, all its data is synchronic, and often also almost bewildering in the richness of its dialect observations. For Evenki, for example, it lists forms from thirty-seven different dialects and other subvarieties of the language, along with fifteen for Lamut, two for Negidal, two for Oroči, four for Udihe, and three for Nanai. Only Solon is represented by but a single vernacular, and that, surprisingly enough, not the result of new or even recent fieldwork, but basing itself instead on Ivanovskij 1894 and Poppe 1931, which monograph appears to be the most recent data available, even in the USSR, on this (perhaps now extinct?) language. All in all, and taking languages, dialects, and subdialect varieties together, TMS provides us with data of greater or lesser detail and completeness for the truly astonishing total of sixty-seven different linguistic entities.

Particularly in view of these numbers, and in consideration also of the evident diversity of the Tungus languages, it is only natural that attempts

toward their classification have hardly been lacking in the literature. For some decades our Soviet colleagues have worked within the system of the so-called NIA INS (Naučno-issledovatel'skaja Asociacija Instituta Narodov Severa) classification, as worked out by linguists in Leningrad in the early 1930s, and apparently codified in Černjakov 1935 (not seen). This is an inherently simple and convenient system, one that divides the ten principal languages numbered and listed immediately above into two major groups in terms of their present-day geographic distribution, resulting in a 'Northern' or 'Tungus Group' consisting of Evenki, Negidal, Solon, and Lamut, and a 'Southern' or 'Manchurian Group,' including Manchu, Nanai, Olča, Oroki, Udihe, and Oroči; from this classificatory scheme also follows the usual modern Russian designation of these languages as the 'Tungus-Manchu' family (as for example in the title of the dictionary that we, following the Soviet colleagues, cite as TMS).

To the limited extent that classificatory considerations play a part in its articles, this system is also the one employed in TMS, and is essentially also the classificatory basis of Benzing 1955, as it was that of Cincius 1949 upon which his monograph so heavily draws. Menges (1968:25-30) offers what is probably the most useful and detailed summary of the details of this classification in a readily available source, especially convenient for its information about the several (admittedly minor) changes and alterations that the system has from time to time undergone during its utilization in the USSR.

This same essentially 'North-South' scheme for the classification of the Tungus languages is also taken as a given in virtually all of the few original works on the Tungus languages by non-Soviet scholars, e.g., Ikegami 1971, an independent reconstruction of proto-Tungus phonology with certain original features, though not totally without debt both to Cincius and Benzing.

In the 1970s, Doerfer turned his attention to the question of the classification of the Tungus languages. His final solution for this problem is conveniently summarized in Doerfer 1978, which also contains references to several of his other earlier papers that present more elaborate documentation of his views and findings. One cannot but agree with Doerfer on the necessity for initially questioning and eventually replacing "the concept of Northern and Southern Tungus" (1978:2). But beyond that, it is difficult to understand, much less to follow unequivocally, all the details of his proposal for applying to questions of Tungus linguistics what he modestly describes as his own "procedure ... ideal for every linguistic classification" (1978:3).

The principal difficulty attending Doerfer's "ideal procedure" is his lack of attention to questions of goal and aim: he never once makes it

clear *why* he is concerned with classifying the Tungus languages, nor does he ever tell us what his classificatory statements are designed to accomplish. Most of the time he appears to be operating on the basis of an un verbalized assumption to the effect that the purpose of his classificatory statements is to uncover some earlier scheme of genetic relationships that underlies the present-day similarities among (and differences between) the languages with which he is concerned — and such covert concern for genetic relationships especially appears to underlie his otherwise undefined employment of the terms ‘related’ and ‘relationship’ more often than not (e.g., 1978:5). But shortly thereafter, and without methodological clarification or definition, he without warning embarks upon issues of what he calls (again, without precise definition) “diachronical and synchronical investigation” (“...a diachronical investigation will often show results differing from those of synchronical research,” 1978:6). Other unclear and undefined terms such as his frequently invoked locution “a special language” additionally cloud the issue, especially when we are first promised that “Jurchen has been found to be a special language,” but then that “[t]he differences between Jurchen and Manchu are minor,” and even finally that “Jurchen is very close to Manchu” (1978:2, 12, 19).

Doerfer’s eventual scheme of classification (1978:5-6) is arranged around a “Northern branch” (La., Ev., Neg., Sol.) a “Central branch” (Oroč., Ud., Na., Olč., Orok.), and a “Southern branch” (Ju., Ma.); the utility of his system is further limited by his consistent reluctance to tell us what his classification is intended to be or to represent, a lack hardly offset by his willingness to tell us what it is *not* (“a general scheme ... identical neither to that of the old-fashioned ‘Stammbaum’ nor to that of the ‘Wellentheorie’...” 1978:3).

Most curiously, and particularly surprising in view of his otherwise extensive bibliographical coverage, Doerfer (1978) took no account of the one earlier contribution to the literature that had not only suggested an important and original revisionist approach to the question of Tungus linguistic classification, but had also — and this virtually without parallel in the literature — addressed the issue of *why* this question is of linguistic significance in the first place. This was the brief note of Ikegami 1969 (published however only in 1974; Doerfer notices this contribution, to be sure, in a note to his paper (1978:26) apparently added in proof, but does not consider the content of Ikegami’s contribution either there or in the body of his own work).

Ikegami 1969, a resume of an apparently unpublished paper presented at the XII PIAC in Berlin, is tantalizingly, even cryptically, brief, but nevertheless it makes two important points: he suggests reworking the

traditional north-south classification for the Tungus languages on the basis of the attested reflexes for phonological entities in the original pTg linguistic unity, and he further suggests “daß diese Klassifikation die Verzweigung des Urtungusischen widerspiegelt” (1969 = 1974:271). Unfortunately the published resume (all that we have of Ikegami’s paper) includes no lexical materials or other specific evidence for his position, and consists only of a general, much abbreviated summation of the phonological isoglosses upon which he proposed to erect his system; unfortunately also even his extremely succinct statements of the details of these isoglosses involve either misunderstandings or serious typographical errors, so that, e.g., his isoglosses 2 and 3 both describe correspondences for “x-,” in effect obscuring the all-important historical-linguistic distinction between pTg *s- > h- and pTg *x- > x- that is essential for any clarification of the genetic relationships connecting most of the attested languages.

Nevertheless, Ikegami’s brief resume tells us rather more than most of the more bulky contributions that preceded and followed it, pointing, as it does, toward the possibility of reworking the traditional north-south classification of these languages for the express purpose of clarifying the earlier relationships, not merely spatial but also historical-linguistic, that have without question existed among the Tungus languages in the course of their long but virtually undocumented history.

Most recently we have also to concern ourselves in this connection with Janhunen 1983, an important if frustrating contribution that is likely to be overlooked because it was published in an obscure journal, and to remain unread by Altaicists because it appeared only in a Japanese translation. Janhunen sets forth a most ambitious scheme for the classification, not only of the languages but also of the races (Jap. *minzoku* ‘Völker’) of all northern Asia. Unique to his paper, and valuable, is the author’s realization of the necessity for making a clear distinction between descriptive-geographic categorizations on the one hand and classification based upon (and hence one would hope also reflecting) historical genetic relationships on the other (esp. 1983:48 sqq.). For Janhunen, the Tungus *Völker* were originally one of a total of eleven totally independent, totally unrelated linguistic communities (Jap. *gengo kyōdōtai* ‘sprachliche Gemeinschaften’), all eleven of which were already in existence in various parts of north Asia by the neolithic, by which period they had all already seen “several thousand years of independent development” (1983:50). Janhunen locates the original Tungus speakers near the eastern shore of Lake Baikal in the neolithic, from where he sees them moving to the east, and subsequently dividing

into what eventually becomes, for him also, in effect the traditional north-south classification.

Despite its refreshingly new methodological approach to the question, Janhunen 1983 can be of little service. The author has totally dispensed with the objective correlative, so that we are never shown or instructed but simply told. No single linguistic form is ever cited in evidence for the historical vicissitudes of any of the hundreds of languages with which he deals; and yet he purports to know so much about their earlier relationships — or with Janhunen what is more often stressed, their early non-relationships — that he is able to draw a set of precise maps of north Asia exactly locating his eleven linguistic communities, including Tungus, all in splendid isolation from one another, as he insists that they must have existed in neolithic times. Given as we are no single clue to how Janhunen arrived at any of these data of which he writes with such conviction, we are at a loss to know what more to do with his claims, despite the novel and welcome methodological orientation of his contribution.

Especially against the background of Janhunen's entirely unsupported speculations, the urgent need for a basic reworking of the traditional classification of the Tungus languages — which, when all is said and done, even Janhunen simply takes over without comment — becomes especially clear, particularly when we attempt to correlate such undocumented claims in terms of what is known about Tungus comparative phonology — most notably if we inspect, e.g., the several sections in Benzing's sketch of pTg phonology. Then it is immediately apparent how decidedly, and also how widely, the traditional 'north-south' classification of these languages, which Benzing 1955 utilized throughout, diverges in fact from the very historical-phonological data with which he is attempting to correlate it.

For a concrete illustration, we have only to glance at a typical entry from Benzing's phonology (1955:989):

tg.	ma.	go.	olč.	orok.	oroč.	udh.	sol.	negil.	ew.	lam.
*x-	o	h	x	x	o	o	o	o	o	o
*zi-	i- (ni-, si-)	si-	si-	si-	i-	i- (si-)	i-	i-	i-	i-
*s-	s- ~ š-	s-	s-	s-	s-	s-	s-	s-	s-	h-
*si-	si-, hi-	si-	si-	si-	si-	si-	si-	si-	si-	hi-
*p-	f-	p-	p-	p-	x-	x-	o	x-	h-	h-
*pi-	fi-	pi-	pi-	pi-	xi-	si-	i-	xi-	hi- (~ i-, si-)	hi-

Here, as throughout, Benzing has arranged the Tungus languages he cites in evidence from 'south' to 'north,' with (using his siglia) "ma., go. [i.e., Na.], olč., orok., oroč., udh." for 'south' and "sol., negd., ev., lam." for 'north.' But the attested reflexes that he records for, in this case, pTg **x*-, whether that phoneme stood before the vowels generally or specifically before pTg **i* or **ī*, give the lie to this same 'north-south' classification, at least as a scheme representing any earlier set of intrafamily linguistic relationships. Within his supposed 'south' group, it is clear that Nanai, Olča, and Oroki are distinguished, both from the other supposed 'south' languages, as well as from the so-called 'north' ones, by virtue of their preserving original **x*- as (with Benzing), *h*, *x*, and *x* resp., as against the \emptyset of Manchu, Oroči, and Udihe. (Actually, Benzing's Na. *h* is a misleading misprint throughout, Nanai having regularly *x* in the words in question, cf. TMS passim; the point is important, since *h* does occur in other Tungus languages, but always as a reflex of earlier **p*- or **s*-, as set forth accurately enough in the same table reproduced above; but *x* in Nanai is always and only the reflex of **x*-.)

Even when, as in this fashion, we clear up this and a few other ambiguities in Benzing's data in the light of our more recent materials (even though this particular problem of the Na. *h* could have been solved by consulting Cincius 1949, Benzing's principal source!), we end up only making the same point all the more strongly, i.e., that this and a number of other correspondences in Tungus historical phonology show that, whatever else may be said in its favor, the traditional 'north-south' classification of the Tungus languages tells us virtually nothing about their original historical relationship to one another, and by the same token, also next to nothing about what their earlier geographic locations vis-à-vis one another may have been.

Much the same indication of serious discrepancy between the traditional postulation of a significant 'north-south' axis of bifurcation among the Tungus languages is similarly apparent from an inspection of the developments in the various languages of certain original medial consonant clusters that must be reconstructed in order to account for the attested reflexes in the linguistic evidence.

The problems, as well as the historical-linguistic explanatory potential, presented by one such typical cluster, pTg **-rb-*, were explored in some detail in Miller 1985 (a somewhat expanded version of a paper first read to the XXV PIAC in Uppsala in 1982). There, in the course of studying the vicissitudes in Altaic in general and in Tungus in particular of a long-standing and widely-spread I.-E. loanword, **ser-p-* 'sickle, hook',¹ it was discovered that the forms in question offer particularly rich materials

for the precise investigation of these issues because of the fortunate accident of the existence of a number of early Chinese transcriptions of several different versions of this I.-E. root as it developed, after the borrowing, within proto- and pre-Tungus. We concluded that, after the initial borrowing of this **ser-p-*, assimilation in voicing soon thereafter resulted in **serb-*, after which the various Tungus languages, as may be demonstrated on the basis of the independent documentary evidence of the Chinese transcriptions, subjected this **-rb-* sequence to three different treatments.

In the case of **serb-*, significant parallels for these three different treatments may be exemplified by the reflexes, in the various languages, of pTg **gerbü* 'name', along the following lines: Ev. *garbi*, La. *garb*, Sol. *garbi*, Na. *garbu*, against Neg. *galbi*, Olč., Orok. *galbu*, and Oroč. *gäbbi*, Ud. *gäggi*, Ma. *gebu*, and Jr. **go-ḥu*. The consideration of these parallels in phonological development makes it clear that the traditional 'north-south' dichotomy does not fully explain the attested linguistic data. Accordingly, it was concluded (Miller 1985 *passim*) that earlier in their history many if not most of these languages must actually have stood in rather a different geographic relationship to one another than the traditional classification scheme suggests. In other words, it was proposed that these specific historical isoglosses for pTg **-rb-* are all in one way or another indicative of earlier spatial alignments within the proto-Tungus linguistic unity, but that these earlier alignments were disturbed or sometimes totally obliterated by subsequent geographic displacements.

Finally, and preceding along the lines sketched above, it was postulated that original pTg **-rb-* was treated as follows: (1) the sequence appeared unaltered and unshifted as *-rb-* in the prototype(s) of what we may now dub an 'Original Northern' group of languages, languages now represented by Evenki, Lamut, Solon, and Nanai, all with their uniform *-rb-*; (2) it appeared partially altered, and with *l* for **r*, as *-lb-* in the prototype(s) of what we may now dub an 'Intrusive' or 'Transitional' group of languages, the prototypes of the languages now represented by Negidal, Olča, and Oroki, with *-lb-*. This intrusive or transitional group was intermediate both in terms of original geographic-spatial relations and also in terms of stages of linguistic development to a third group, where (3) it appeared, largely altered, with assimilation of **r* (or **l*) to the following stop, either partially, as *-gb-*, or totally, as *-bb-*, and this further simplified to *-b-*, in the prototype(s) of what may now be dubbed the 'Original Southern' group of languages, the prototype(s) of the languages now represented by Manchu, Jurchen, Udihe, and Oroči. But in the case of this 'Original Southern' group also,

a further subdivision is possible, (4) in terms of the extent to which the consonantal assimilation affecting the cluster advanced, thus (4^a), 'Early Original Southern,' with the *-gb-* of Udihe being the least altered, and with Oroči's *-bb-* being somewhat more changed, alongside (4^b) 'Late Original Southern,' with simplified *-b-* in Manchu and Jurchen representing the result of the final and most sweeping set of changes in the development as a whole.

Central to this hypothesis of Miller 1985, and to the present contribution as well, is an important postulate of historical-linguistic methodology, a principle frequently to be found implicit in dialect-classification studies, but one that is by that same token only rarely, if at all, overtly verbalized.

Most briefly stated, this central assumption may be described as the logical inversion of a historical-linguistic principle first enunciated just over a century ago by Jos. Schmidt, to the effect that "special resemblances can be found for any two branches of I.-E., and ... these special resemblances are most numerous in the case of branches which lie geographically nearest each other" (Schmidt 1872, paraphrased in Bloomfield 1933:317, §18.12); this was, of course, Schmidt's classic treatment of the problem of overlapping features of special resemblance among the I.-E. languages that was eventually formalized under the so-called 'wave hypothesis.'

The methodological inversion of this principle herein proposed seeks to substitute cause for effect. If, as indeed appears to be the case, special resemblances between any two branches of a language family such as I.-E. are most numerous in the case of branches presently to be found geographically nearest each other, it is consequently suggested that other instances of special resemblances between two or more related languages or groups of languages that today do not lie geographically near one another are most readily to be explained by their having earlier been so located — or as we have already (Miller 1985) suggested, in applying this same concept with particular reference to Tungus, when the traditional 'north-south' classification obviously does not explain the distribution of attested reflexes for either unitary or complex phonological units, we can only conclude that earlier in their history many of these languages must have existed in different spatial relationships to one another than the traditional classification now suggests. Finally, we also make, and in the present contribution shall attempt to exploit, the assumption that, just as reversing the logical basis of the principle of regular sound change ('phonemes change') makes possible the reconstruction ('recovery') of earlier phonological entities and their systems, so also does reversing the logic of Schmidt's hypothesis make possible the reconstruction of the

earlier geographic, in the special sense of the earlier spatial, relationships of languages and branches of linguistic families showing special resemblances.

Presumably no apology need be offered here for the overall assumption, implicit throughout the above, that methods and principles of historical linguistics originally devised from I.-E. studies may indeed have a simple and direct application to questions of Altaic linguistics. Many will surely argue against this; some already have (e.g., Doerfer 1974:108: "Die idg. Wellentheorie ins Spiel zu führen, wäre gewiß verfehlt ... Im Grunde weisen die altaischen Sprachen ganz verschiedene und unvereinbare Wortschätze auf, die nur mit recht losen Fäden verknüpft sind"). But to many others of us the point seems beyond fruitful debate, all the more so because the utility of this I.-E.-derived method for bringing into historical-linguistic order "incommensurable forms and partial similarities that cannot be reconciled with the family-tree diagram" (Bloomfield 1933:318) is obviously made-to-order for treating many of the most striking problems in comparative grammar that the Altaicist encounters. Surely I.-E. studies have no monopoly on the observation that no speech community is ever quite or absolutely uniform, no more than on the equally important observation that while, on the one hand, history sometimes shows us a sudden cleavage, or family-tree-type of change, it also frequently shows the division of a speech community by splitting, typically due to the intrusion of a foreign community, and equally often the evidence of irreconcilably different forms, whose relation to one another, whether as alternants or as dialect variants, it is unable to clarify (Bloomfield 1933:311-19, §§18.10-18.13). From this it is only a short step to Schmidt's solution, and to the wave-hypothesis, which is, after all, merely another way of describing the disappearance of intermediate dialects. Our position is simply the essentially Darwinian hypothesis that like effects operative in like circumstances may be expected to yield like results; once we admit the utility of the comparative method of the neogrammarians for the study of any earlier non-I.-E. linguistic unity, we can then hardly arbitrarily exclude this or that language group from the overall implications of that method.

Doerfer mentions (but does not illustrate in specific detail) the need for "consider[ing] as many characteristics as possible (e.g., even 100)" (1978:3). In this initial exploration of the ramifications of the hypothesis set forth immediately above, with immediate application to the history of the Tungus languages, we shall be somewhat less ambitious, and deal in the main with the earlier geographic implications of the evidence in the attested languages for two proto-Tungus consonants in morpheme-initial

position, pTg **p-* and pTg **x-*. Each of the two represents an important, historically significant feature of the phonology of the original linguistic unity; each also requires in turn a number of brief explanatory statements.

By pTg **p-* we mean the original phoneme represented in the attested languages by the following set of correspondences: Ma. *f-*, Na. *p-*, Olč. *p-*, Orok. *p-*, Oroč. *x-*, Ud. *x-* (but *s-* before **i*, **ī*), Sol. *∅*, Neg. *x-*, Ev. *h-*, Lam. *h-*. In larger Altaic terms, this pTg **p-* corresponds to pA **p-*, itself the correspondence of Middle Mongol (MMo) *h-*, Monguor *f-*, (with regular alternations as *x-* and other positional variants), Turkic *∅*, but with certain south-east and south-west Turkic languages showing the survival of a proto-Turkic *h-* (Poppe 1960:10, §7), Korean *p-*, O(ld) J(apanese) *F-* (> N(ew) J(apanese) *h-* ~ *f-* as regular positional variants) (Miller 1971:305, nos. 2-5, where most of the gaps there indicated may today be supplied).

Four examples of the correspondences for pA **p-* must here suffice: (1) **wart**:² Mo. *egüün* < **pögü-n* 'wart', Khalka *ū* 'id.', Ev. *heŋte*, La. *heŋtá*, Neg. *pučiekte* 'pimple', Olč. *puktu/e*, Orok. *peukte*, Na. *pükte*, Ma. *fuhu* 'wart', M(iddle) J(apanese) *focuro* 'mole (on the face)' (cf. Poppe 1960:11, 61, 133, except for the MJ comparison); (2) **tremble**: pA **pürk-*, MMo. *hürgü-* 'become shy, get angry', OTurkish *ürk-* 'be frightened', Olč. *pučujuli* 'vexatious', Orok. *pusuipu-* 'endure', Na. *počopočo* 'with annoyance', Ma. *fuče-* 'get angry, mad', late OJ *FuruF-* 'shake, tremble', secondary formation to OJ *Fur-* 'move (tr. and intr.); summon spirits' (cf. Poppe 1960:80, 83 bis, 87, 111, and TMS 2.45b); (3) **shaft**: pA **pāra/īl₂*, Mo. *aral* 'Deichselstangen, Karrenboden', MMo. *haral* 'id.', Kirghiz, Tatar, Toból, *arış* 'Deichselstangen', Ev. *hara* 'wedge; pommel (reindeer saddle)', Olč. *para* 'sleigh', Orok. *pawūra*, *paura* 'shaft', Na. *para*, *fara* 'sleigh', Ma. *fara* 'sleigh; edges or shafts of a wagon', Korean *palko* 'sledge', NJ *hari* 'beam, girder, crossbeam' (Poppe 1960:11, 40, 77, 78-9, 122; TMS 2.316b-317a); (4) **long**: pA **pur₂-*, Mongol *urtu*, MMo. *hurtu* 'long', Monguor *fudur* 'id.', OTk. *uzun*, Chuv. *vêrêm* 'id.', OJ *Furu-* 'long (of time)' (Poppe 1960:81). The correspondence is one of several that are central to the hypothesis of an earlier Altaic linguistic unity; it would hardly be an overstatement to designate its original formulation by G. J. Ramstedt in 1916 as the birthday of comparative Altaic linguistics (or perhaps even earlier, by P. Schmidt in 1898, as suggested in Boodberg 1939:298 note 15?).

Any attempt at a succinct account of the fate of pTg **x-*, and also of that of the pA **x-* of which it was the direct continuation, involves somewhat greater complications. Poppe's 1960 reconstruction of Proto-Altaic phonology made no provision for this phoneme. Following the lead of Ramstedt, Poppe argued that the many cases of Tungus

lexical cognates in which \emptyset in most of the languages, corresponding to \emptyset in Turkic and Mongolian, also corresponded to Na., Olč., and Orok. x ($\sim s$ before $*i$, $*i$) did not offer evidence for an original unit in the proto-language's phonological system: 'Es ist kein Grund vorhanden, ein ursprüngliches $*h$ - oder $*\chi$ - zu postulieren. Dieser Laut ist nämlich sekundärer Natur' (Poppe 1960:32, §21). The point was, to be sure, still debatable in 1960, but only because Japanese lexical evidence had then not yet been introduced into the question. Once that was done, it became clear that we must indeed reckon here (in one of our extremely rare disagreements with Poppe's 1960 reconstruction) with an original phoneme that was lost ($> \emptyset$) in most, but significantly not in all, the so-called 'inner Altaic languages,' but which survived to leave substantial amounts of evidence in portions of Tungus, and also in Japanese. A detailed discussion of this important innovation, now necessary in the classical Ramstedt-Poppe reconstruction of Proto-Altaic, which contains *inter alia* a substantial number of newly proposed Japanese cognates for words containing original pA (and pTg) $*\chi$ - is now in hand (Miller 1987). This contribution demonstrates why the involved nature of the Japanese evidence for pA $*\chi$ -, in particular the complex pattern of the phonologically determined occurrence of the Japanese reflexes for this original phoneme vis-à-vis the original Proto-Tungus vocalism, makes it necessary to reconstruct this element for the original language, and to abandon the Ramstedt-Poppe hypothesis that sought to explain the Tungus reflexes for the same as the sporadic appearance of a secondary element of prothetic articulation. Thus, for pA and pTg $*\chi$ -, OJ has a regular pattern that involves the ordered distribution of three different Japanese phonemes as attested reflexes for the postulated $*\chi$ -, but with each of these three phonologically-structurally determined according to the quality of the vowel that originally followed the $*\chi$ - in the Proto-Tungus linguistic unity.

The three Japanese phonemes involved in this pattern of distribution are OJ F , OJ w , and OJ s ; before the OJ reflexes of pTg $*\check{a}$, $*\check{u}$, $*\check{o}$, and $*\check{\check{o}}$, but OJ w as representing $*\chi$ - before $*\check{a}$ and $*\check{u}$, and OJ s as representing $*\chi$ - before $*\check{i}$ and $*\check{\check{i}}$. This is exactly the kind of rigorously determined scheme of reflexes that can hardly be the result either of chance or of borrowing, and that as a consequence speaks vigorously on behalf of ultimate genetic relationship.

It will be sufficient here to give the following eight representative examples in order to demonstrate the variety of lexical evidence now available for this hypothesis (selected from among a total of 36 new etymologies published for the first time, and in greater detail, in Miller 1987): (1) **dig out**: pA $*\check{x}\check{a}r_2$ - > pTg $*\check{x}\check{a}r$ -, Ev. *er*- 'to dig (snow), excavate

(earth)', Orok. *xeri-* 'to dig (snow)', Olč. *xeru-*, Neg. *xeruči-* 'id.', Ma. *eri-* 'to sweep (a floor)', TMS 2.46a-b, Written Mongol (WMO.) *erü-* 'dig, cut, hack', Chaghatai, Osman *āz-* 'rub, grate', Poppe 1960:103, OJ *wer-* 'engrave, incise (on wood, stone); dig out (a hole in the ground)'; (2) **bark**: pTg **xūra(.kta)*, Ev. *ūra-* 'cover (of a tent, etc.)', *ūran* 'summer tent of skins or bark', Orok. *xūrakta*, Olč. *xorakta*, Na. *xorakta* 'bark', Ma. *umriha*, *uriha* 'membrane, inner bark on trees, skin on walnuts and hazelnuts', TMS 2.282a, early MJ *Foro* 'supplementary armor worn on the back by a mounted warrior to ward off arrows', WMO. *urača* 'Laubhütte, konische Hütte aus Birkenbast, Art Wigwam', Yakut *urasa*, *uraha* 'id.', Poppe 1960:80; (3) **finger**: pTg **xunia*, Ev. *un'akan*, *un'akāsān* 'index finger', Orok. *xūná* 'spoon', *xūn'aptū(n-)* 'finger-ring', Olč. *xon'a(n-)* 'spoon', Na. *xoñā* 'spoon', *xoñaka* 'finger-ring', TMS 2.276a-277b, OJ *Fone* 'bone' (cf. the semantics of MLG *knokel* 'bone, knuckle', against Danish *knokkel* 'bone', NHG *Bein* 'leg', cognate with OE *bān*, NE *bone*); (4) **enjoy**: pTg **xāpī-*, Ev. *ewī-*, Orok. *xupu-*, Olč. *xupī-*, Na. *xupī-*, Ma. *efi-* 'play, enjoy oneself; act', OJ *weF-* 'be intoxicated', OJ *werak-* 'laugh aloud with pleasure (esp. when drunk)'; (5) **flower, ornament**: pT **xīlaga*, Ev. *ilaga*, Orok. *silla*, Olč. *irga* 'ornament', *sila* 'flower', Na. *elga* 'ornament', *sela* 'flower', *sila-* 'to bloom', Ma. *ilha*, *ilga* 'flower, blossom; patterned, colored', Ma. *ilgari*, *ilhari* 'paper strips attached to a willow branch used as an offering to spirits', TMS 1.304a-b, OJ *siraka* 'ornament or device (of obscure construction but involving paper or other white material), employed in the worship of autochthonous deities'; (6) **snow**: pTg **xīman.sa*, Ev. *īmanna* 'snowfall', Orok. *sīmata* 'snow', Olč. *sīmata*, Ma. *nimangi* 'id.', TMS 1.312b-313a, OJ *simo* 'frost', Ryūkyū *simu* 'cold winter rain'; (7) **tail**: pTg **xūrgū*, Ev. *irgi*, Orok. *xudu*, Olč. *xūjū*, Na. *xujgu* 'tail', TMS 1.325a-b, OJ *worō* 'tail (of mountain bird)', OJ *worōti* 'the great serpent (in the myth)', cf. Ev. *irgiči* 'wolf (lit., 'the tailed one')' Korean *ilhūi* 'wolf'; (8) **star**: pTg **xōsī(.kta)*, Ev. *ōsīkta*, Orok. *wasīkta*, Olč. *xosta*, Na. *xosekta*, Ma. *usiha* 'star', TMS 2.27a-b, OJ *Fosi* 'star' (and as a consequence not cognate with its frequently-cited semantic and phonetic 'look-alike' New Korean *pyöl* 'star', which goes instead with pTg **pālā(n)* 'elk', Na. *peule* 'Ursa Major', Ev. *heglun*, *hoglan*, *heglun*, *evlen* 'id.', lit. 'young of the elk', TMS 2.306a).

With this new historical-comparative evidence from Tungus, and Altaic, phonology in hand, we are finally in a position to undertake the major enterprise of the present contribution, i.e., implementing our proposed reversal of Schmidt's dictum in such a manner as to make it feasible to estimate the original (or at the very least, the considerably earlier) geographic relationships of the Tungus languages with respect to one another, on the basis of the special resemblances that they share,

assuming that this sharing is best (i.e., most elegantly and economically) to be explained by prior geographic, resp. spatial proximity sometime during their earlier linguistic unity.

To carry this out, we once more have resort to much the same variety of historical-linguistic mapping that we first attempted, in very general terms, for the entire Altaic domain in Miller 1971:46 Fig. 2. In that figure we aimed "to give some graphic indication to the probable loci of earlier linguistic unities within the larger proto-Northern and Peninsular Altaic area ... [with particular attention to] the geographical area of linguistic symbiosis ... that must have been shared in common, during the proto-Northern and Peninsular Altaic period, by speakers of earlier forms of Tungus on the one hand and those of earlier forms of Korean-Japanese on the other" (1971:45). On the present occasion we continue and elaborate upon that same concern for original geographic relationships and attempt to determine what earlier spatial scheme most succinctly explains the later attested reflexes, not only in Tungus, but also in Altaic as a whole, for pA **p-* and pA **x-*. In contrast to the Old Japanese reflexes of **x-*, Korean (at least for the time being) has revealed no significant lexical evidence within this sector of the phonology; hence our present concern is necessarily limited to Old Japanese, in addition to Tungus and the other traditionally so-called 'inner Altaic materials,' i.e., Turkic and Mongolian.

This schematic diagram or putative historical-linguistic mapping — either term is to be preferred to the potentially misleading 'map' — largely speaks for itself, and requires but little comment concerning a few matters of detail.

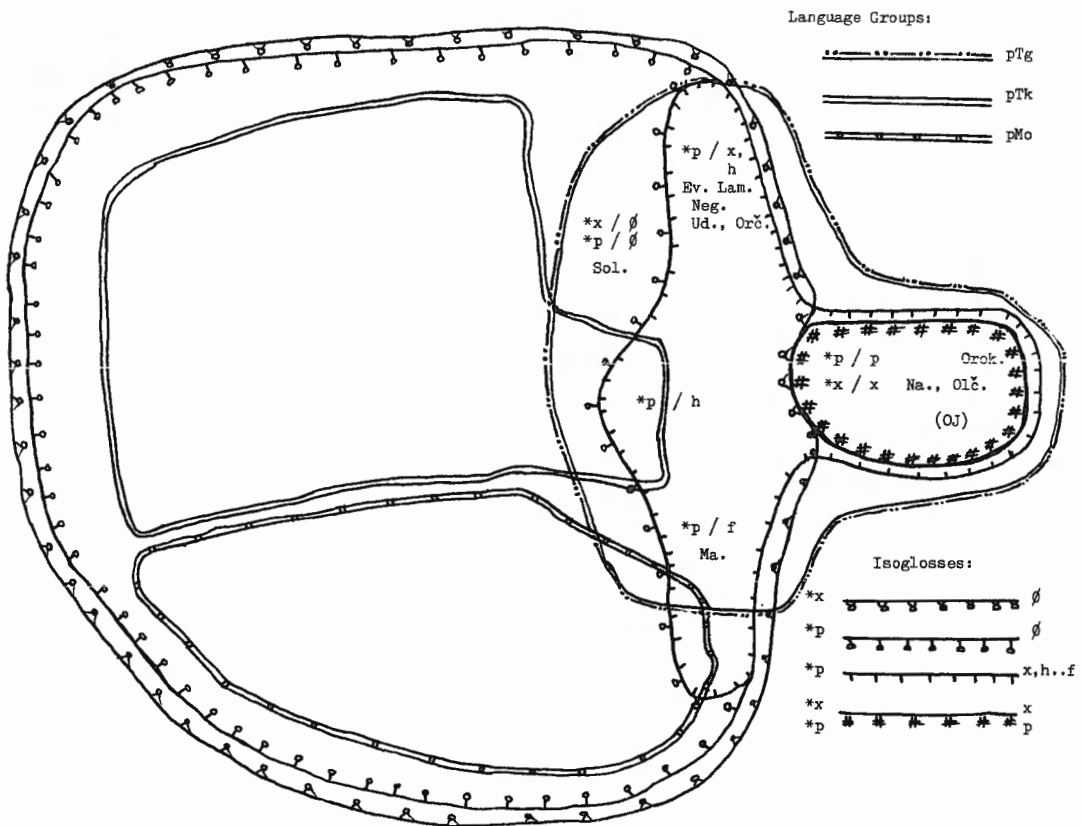


Figure 1

In principle, it seeks to tamper with present and/or attested geographic locations as little as possible, adjusting these sites only as necessary in order to bring the attested phonological correspondences into meaningful (i.e., spatially adjacent, hence shared) relationships. The overall, most extensive perimeter of the diagram represents the greatest expanse, and hence also the probably overall configuration, of the Altaic linguistic unity shortly before its initial fragmentation (the first of the two major 'Altaic homelands,' on which see Miller 1980:54-55). Proto-Turkic and Proto-Mongolian are to be found more or less in their traditional and expected locations, in the upper left (i.e., northwest) and lower left and center (i.e., southwest) of the original linguistic domain. Proto-Tungus is best explained, and accordingly is plotted in the diagram, as originally occupying the right (i.e., the eastern) sector of the domain. Within this sector, however, the major earliest geographic demarcation evidently was not the familiar 'north-south' dichotomy of the traditional Tungus classification schemes, but a somewhat more complex distribution in which a nodule-like eastern-reaching extension of the Tungus sector developed early in the history of this family, providing in its turn the nucleus for later developments largely responsible for Orochi, Nanai, and Olč'a, as well as for significant elements later involved in the genesis of Old Japanese.

Furthermore, the much-commented upon and familiar 'south' orientation of Manchu, while still of obvious importance in this new scheme of geographic relationships, is now of more interest for the manner in which it relates the *f* reflexes of **p-* in Manchu with the *f*, *x*, and other reflexes for this same phoneme in Monguor, and also for the way in which this relative positioning of the Manchu evidence vis-à-vis Proto-Mongolian furthermore provides an earlier common ground of shared phonological developments that in turn account for MMo. *h-* < **p-*, in contrast to the general (but later) Mo. \emptyset < **p-*, which development is however here also accounted for by the provision of a contiguous area for common developments shared by Turkic and Mongolian, additionally reinforced and validated by the provision of a distinct region for Solon, with its \emptyset for both **p-* and **x-*.

Perhaps the most significant element in this schematic diagram, particularly for our understanding of the relationships between Japanese and Altaic, is the nodule to the right, i.e., to the east, where both **p-* and **x-* were retained; this nodule can hardly be other than the graphic-spatial recovery ('reconstruction') of the earliest impetus toward the eventual breakup of the original Altaic linguistic domain, and as such an event of major historical, as well as of linguistic, significance for any account of what eventually happened to early man in Greater Asia.

As was true also in the case of earlier attempts at this same variety of historical-linguistic mapping (e.g., Miller 1971:46), it is probably equally necessary to emphasize what the figure in question is **not** intended to represent as it is to explain what it can, and does, tell us. Most importantly, it is **not** a road map, particularly not a road map in the sense that the figure in Janhunen 1983 attempts to be one, i.e., a literal map overtly oriented onto the actual physical map of Greater Asia. Nor, it goes virtually without saying, does it aim either at the detail or at the coverage of the map entitled 'Die tungusischen Völker und ihre Dialekte' on the unnumbered page enclosed with Doerfer 1978, listing as it does hundreds of languages and dialects in their presumptive modern locations, but simultaneously attempting to exhibit information about their earlier linguistic connections with one another (in part by an elaborate system of color-coding that has in several important particulars baffled the printers, and as a result has left the reader also at something of a loss). Returning to essentials (and to the beginnings of our science as well), Figure 1 above is nothing more or less than a Tungus-oriented experiment with a mapping on the model of that figure reproduced long ago by Bloomfield (1933:316, Fig. 3): "Some overlapping features of special resemblance among the Indo-European languages, conflicting with the family-tree diagram. — Adapted from Schrader." Instead of committing itself precisely and unequivocally to the question of specific, definite geographical sites, it rather attempts to exhibit historical-linguistic data solely in terms of those relative spatial alignments, the postulation of which appear to offer the most economical explanation for the later interrelationships of these languages, as subsequently attested in the form of the reflexes participating in the sound correspondences upon the basis of which we are able, in the first place, to postulate all linguistic relationships earlier existing in this area, Altaic as well as Tungus.

In this respect in particular, the figure above, as a result, closely parallels, both in purpose and in plan, as well as in its principal premise, the 'Fig. 2, Proto-Northern and Peninsular Areal Relationships' (Miller 1971:46) to which we have already made reference above. As in that earlier mapping, now also we propose that, since the traditional graphic display along Stammbaum lines had necessarily to pass over, if not actually to obscure, a considerable number of "essentially areal relationships — or more exactly, perhaps, historical relationships that can best be explained in areal terms," it would be more useful to attempt to provide "some graphic indication to the probable loci of earlier linguistic unities within the larger proto-Northern and Peninsular Altaic area" (Miller 1971:45) — which is precisely what we have attempted to do with

the figure above, no more no less. The sole additional methodological refinement of the figure above, in contrast with that published in 1971, concerns the stipulation that the present mapping may most usefully be viewed as representing the relative spatial relationships of the proto-languages as they appeared sometime early in the history of proto-Altaic, but not at the earliest stage of that linguistic unity that we may now recover, rather at some point shortly after the initial migration away from the first Altaic homeland, but probably also shortly prior to their arrival in the second homeland (cf. the discussion of these two loci in Miller 1980:54-56 and *passim*).

Virtually — but not entirely — needless to add is the repetition of a reservation apparently always necessary in all postulations such as these, to the effect that, only naturally, what we find in any study directed toward a graphic display such as that in Figure 1 above is the direct outcome of what we set out to look for in the first place. Another way in which to express this same important fact is to remind ourselves yet once more that the scheme displayed above has specifically been focused upon a few specific features from within a considerably larger inventory of items available in the historical phonology of the Tungus languages, particularly and especially because the features chosen here now appear to throw particularly valuable light upon the Altaic-Japanese relationship, and hence stand potentially to enhance our understanding of that important set of still mostly unexplored linguistic relationships in Greater Asia, at the same time that they may properly be expected to throw light upon a number of important earlier interrelationships within the Tungus languages themselves. As a consequence, what is now most urgently called for in future investigations of these questions is the substantial expansion of the parameters of the scenario exhibited in the figure above, through the factoring-in of additional isoglosses: for each such addition, it is reasonable to expect a consequent clarification, perhaps often of an exponential order of explanatory power, in our overall understanding of precisely what constituted the many and involved events that at one time or another transpired in this specific sector of human linguistic history.

In approaching the question of such additions, it will incidentally be of the utmost importance, from the outset, to make the attempt to distinguish between quantity and quality. As is always true in all aspects of historical linguistics, the important issue is not the sheer numbers of our examples, but their historical-explanatory powers, in which sense a small number of “quality” examples always outweighs in importance a large number of trivial or anecdotal cases. At the same time, all possible care must be taken to ensure that the isoglosses selected for future scrutiny are factored in not simply at random, but in a carefully

considered **order** that will, as far as possible, reflect their own grammatical-structural identities. To speak, much less even to think, with Doerfer of "as many characteristics as possible (e.g., even 100)," then to count the agreeing and the deviating features and to enter the contrasting ones onto a general scheme" (1978:3) is plainly to confound the concept of quantity with that of quality; perhaps it cannot be repeated often enough that in this, as in all historical-linguistic work, what matters is not how many examples we have but what kind they are.

In this present preliminary contribution, we have begun with selections from certain critical sectors of the phonology, and this for a number of fairly obvious reasons. Phonology must be, in any ordered system of linguistic analysis, the starting place for all operations. At the same time, those specific sectors of the phonology here selected (especially **p-* and **x-*) are such as might reasonably be expected to supply, even in a purely preliminary investigation such as this, a maximum of insight into what happened in history, because of their key structural role in what we presently understand of the phonological system of the Altaic linguistic unity. As later investigations advance these inquiries further and further into other sectors of the phonology, particular attention will necessarily be due to, e.g., the so-called "*r*-Verschleiss" generally identified as a typical "south Tungus" phenomenon (Menges 1968:205), as well as to a number of other important but as yet still imperfectly described (and hence incompletely understood) combinatory developments of unitary phonemes, e.g., pTg **xo-* > *wa-* and the like (Menges 1975:50, 65, 93 ff.). Thereafter, and on a still higher level of the phonology, there also remains for consideration the still virtually unexplained question of the presence or absence (i.e., of the preservation or loss) of varying degrees of "vowel harmony" — or more properly, "sound harmony," since the phenomenon essentially affects consonants as well as vowels, in the various Tungus languages. At present it would appear that Evenki, Negidal and Lamut virtually or entirely lack this feature, while Nanai, Udihe, Oroki and Olča show it to a greater or lesser extent, perhaps in direct proportion to their present northern-southern geographical distribution; but on the other hand, Manchu, the prototype of the traditional southern group, maintains only vestigial traces. This entire question is an especially critical issue for our eventual understanding of the relationship of Korean to Tungus, if only because of the well-known and extraordinarily regular sound-harmony system of Middle Korean (where the phenomenon affects only the vowels), a feature that we may hardly expect to understand historically without at least a glance at the less systematic and generally skewed systems in the nearby languages.

Finally, at the apex of the phonological node, there remains the tremendous problem of the reconstruction and historical explanation of the original Tungus pitch (or accent) system. Benzing not only postulated "[e]in starker Unterschied zwischen der Betonung in den stg. Sprachen und der der ntg. Sprachen," but he went the bold step further to assert that this distinction "muß in eine verhältnismäßig frühe Zeit zurückgehen" (1955:999, §62). We have elsewhere suggested that specific traces of this same proto-Tungus pitch distinction may also be identified within Old Japanese linguistic materials, where what was (in Benzing's terms at least) an original 'northern-southern' suprasegmental isogloss is to be seen as reflected in an early 'eastern-western' isogloss (Miller 1981:876-77). The parallels between Japanese and Tungus that we cited in this connection are certainly real enough, and obviously relevant to the issues raised by Benzing in connection with his reconstruction of the original Tungus "Betonung"; but what now clearly requires fresh reinvestigation is the relevance of an early 'northern-southern' Tungus suprasegmental dichotomy to attested features on this level of the phonology. If, as now seems to be the case, the oldest geographical-spatial skewing that we may recover for the Tungus linguistic unity was not actually, or solely, one along a north-south axis but rather one early and deeply involved with an additional east-west dimension, then the Old Japanese evidence, with its own 'east-west' connotations, must surely be viewed in a new and even more important light.

Also necessary to keep in mind in this connection is the curious but easily overlooked detail that Benzing's reconstruction of the proto-Tungus pitch is truly a feat of historical-linguistic legerdemain, since as a matter of actual fact all suprasegmental data relative to tone, pitch, stress, or what-you-will, are totally lacking for all the languages entering into his (and Cincius's) reconstruction. His postulation of an original pitch system for proto-Tungus is solely a hypothetical construct evolved, not to account for later, specific, attested suprasegmental reflexes, but instead for the loss or preservation of segmentals — and sometimes even for the loss or preservation not of attested but of reconstructed segmentals — something that, the more we consider it, the less it enhances the credibility of this particular sector of his reconstruction. Finally, it was onto the anything-but-substantial framework of this postulation that in turn Benzing grafted the concept of an entirely early and original 'north-south' dichotomy within the Tungus languages, something that he took without question as an immutable 'given' in all questions of the phonology and history of these languages. The problems that beset such an approach, particularly in view of the findings of the present

contribution, hardly need be elaborated upon. None of these issues necessarily should be taken to mean that Benzing's reconstruction of the proto-Tungus 'Betonung' is either wrong or misleading in its broad outlines; but at the same time they surely do indicate that these questions, posed as they are at the apex of the phonological pyramid of linguistic structuring, deserve independent and fresh investigation free from the fetters of the traditional inherited 'north-south' postulate within whose limits Benzing did his work.

After phonology, there must next come in their proper order morphology followed by lexicon; the implications of both for our question remain virtually unexplored. A significant number of the suffixes established by Benzing in his reconstruction of proto-Tungus morphology appear to be limited in the distribution of their reflexes in the attested languages in such a manner as ostensibly to belong either to the traditional 'northern' or 'southern' sector of the original linguistic unity; but generally even these show a certain amount of skewing even within terms of Benzing's inherited system of spatial presumptions, e.g., his 'Kollektivsuffix (?) *+g,' of which he writes that it "ist nur in den ntg. Sprachen sicher nachweisbar, im Süden scheint das Go. [i.e. Na.] einige Reste bewahrt zu haben" (1955:1016, §78), or his 'Kollektivum *+sa.g,' "das anscheinend nur in den ntg. Sprachen vorliegt" (1955:1019, §79f), or the presumptive geographic-spatial dichotomy in the morphology of the original case-suffix system, particularly (following Cincius) the concept that "nur 4 Kasussuffixe (Akkusativ, Dativ, Lokativ, Direktiv) allen tg. Sprachen gemeinsam seien," while "3 Kasus (Komitativ, Direktiv-Lokativ, Direktiv-Prolativ) seien nur im Ntg. vorhanden..." (1955:1026, §88).

So also with a multiplicity of problems in the lexicon as a whole. The so-called southern Tungus languages widely preserve an original word for 'small river,' Oroč. *uñi* 'id.', with similar forms in Orok. and Na., cognate with Ud. *uni* 'name of a tributary of the Amur', and with Olč. *uñā* 'origin; source, spring', and *uñū* 'small river.' There appears, however, to be no Manchu cognate for these forms, or indeed any forms in the other Tungus languages; but Old Japanese has preserved a cognate in its word *una(-)* 'sea', generally found as a combining form in compounds such as OJ *unaFara* 'high seas', while at the same time many of the Old Japanese compounds incorporating this form still retained the original 'small river' meaning, thus OJ *unakudari* 'down stream', with *kudari*, a deverbal noun in *-i* on *kudar-* 'travel, come down.' The temptation to reconstruct this word as 'proto-Southern Tg.' **uñi*, and to find significance for its apparent localization within the southern portion of the original Tungus domain in terms of the Japanese evidence, particularly in view of the light that this would obviously throw on the spatial location of at least some of

the migrations from the continent to Japan, simultaneously needs to be weighed in terms of the present contribution's approach (cf. also Miller 1986:45 on this term, particularly in the light of the migrations to the archipelago). In this and many other allied cases, studies of forms only partially attested for the entire Tungus domain will no doubt be rendered both more involved and at the same time more rewarding particularly when, as with **uni*, they have specific semantic reference to identifiable features of the terrain that eventually may even be possible to be correlated with discrete geographical sites.

Finally, there remain the major questions still posed by the large cluster of linguistic features, ranging from phonology through morphology and on into lexicon, in which Manchu stands markedly apart from all the other Tungus languages, traditional 'northern' and 'southern' alike, in a word, that set of divergencies that have led the Soviet colleagues to adopt, and cling to, their term of the 'Tungus-Manchu Languages' for the entire family, even though their actual work with these languages now more frequently than not tacitly ignores the historical-classificatory implications of this designation. The existing Tungusological literature, full of disputation and marked by lack of consensus as it is on virtually every other point and issue, uniformly attributes most of these specific Manchu anomalies to 'Chinese influence' (e.g., Benzing 1955:1024, §87; Menges 1968b:141, to cite only two representative samples). It would obviously be folly to deny the presence and the importance of the 'continuous Sinitic attrition' of which Menges writes for the history of Manchu; but what now must be undertaken is the attempt to sort out those elements of this influence that clearly go back only to relatively recent historic time from among those others that must have been operative in the much more remote past, and then in turn to establish with greater precision the geographic-spatial coordinates of each of these tangled strands in the history of the Sinitic influences operative upon Manchu, and in one way or another also, upon certain others of the Tungus languages as well. Manchu-Chinese bilingualism of the intimate sociolinguistic variety, and on the widespread scale likely to have been necessary in order to account for the "strukturelle Zerstörung durch fremde Einflüsse" of which Benzing (1955:960 §7) writes would, in the light of present knowledge at least, appear most likely only to have been a feature of fairly recent history, notably subsequent to the Manchu conquest of China and the founding of the Manchu-Ch'ing house in 1644 of our era. But the reconstruction of the various sectors of the Tungus linguistic unity inevitably takes us immeasurably further back into history than such very recent events. Nevertheless, the further back we go in Tungus, the more prominent become a number of important linguistic

discrepancies and anomalies separating Manchu from the other languages — and many of these are necessarily of an order of antiquity that would appear to cast serious doubts upon (if not totally rule out) any question of explaining them in terms of Chinese influence. At the same time, as soon as we admit Japanese and Korean into our system of overall, larger Altaic relationships, the problem merely grows in intensity. “Mit Ausnahme des Ma. besitzen alle tg. Sprachen Possessivsuffixe, die offenbar nichts anderes als enklitische Formen der Personalpronomina darstellen” (Benzing 1955:1058, §124) — but precisely the same exception must also be made for Japanese and Korean, where the absence of this feature can hardly be attributed to “foreign influences,” at least not in the particular sense in which this expression is generally employed in the Tungusological literature when attempting to attribute such features to ‘strukturelle Zerstörung’ of Sinitic origins — and particularly not in the case of Japanese, given the plain fact of its insular isolation for virtually all of its early history.

The ultimate causes, or motivations, of the very early regional, resp. dialectal variations that the comparative method now enables us to recover for various sectors of the proto-Tungus linguistic unity may likely — in common with the ultimate reasons for all linguistic change — never be precisely identified; but at the very least, and for the time being, to treat them in terms of the fluidly mobile spatial give-and-take that appears, from these present studies, to have characterized the intralingual areal relations of the Tungus domain, precisely as it did the Indo-European domain, is probably to get closer to a true picture of what actually happened here in history, rather than to continue to attempt to account for changes that must have transpired millennia ago simply in terms of political events and changes of the past three or four centuries. *Mutatis mutandis*, what appears to be true of these frequently-invoked ‘Chinese influences’ would also seem to hold for any secondary linguistic influence of the Mongolian branch of Altaic upon Tungus, even though here the time frame is slightly different, and the possibilities for significant linguistic influences having operated at a significantly remote period in time is somewhat, even though probably not impressively, greater. But in the case of both Chinese and Mongolian, their too frequent citation as an all-purpose explanation for what happened in the history of the divergent Tungus languages, particularly in connection with Manchu, certainly conceals more than it reveals about what actually happened in history; and what actually happened in history is, after all, our principal concern so long as we engage in the study of historical linguistics.

It would be premature as well as presumptive to speak now of conclusions; the study undertaken here is too tentative in many of its details, and too abbreviated in its scope, to admit of any such claims. Nevertheless, this still by the same token does not free us from the responsibility of concluding our contribution with at least a sketch of the major points that, it appears to us, these considerations have engendered, to which outline of salient issues we now turn:

1. The Tungus languages arose as, and emerged out of proto-Altaic in the form or shape of a relatively small, relatively unified, but already geographic-dialectically diverse set of languages, still bearing obvious signs of their original close genetic relationship to one another.

2. Early in the course of this emergence, the Tungus languages polarized into a significant degree of north-south geographic demarcation; but this north-south axis in their spatial relationships, while remaining important in their dialect history down to the present time, has never had the overall importance — nor in most details even the specific weight — generally attributed to it in most of the literature.

3. Even earlier than this north-south polarization, we must also, in the history of these languages, reckon with a somewhat more significant and more essentially basic east-west polarity which erupted probably as a factor of the original emigratory élan that early propelled the Tungus speakers out of and away from the proto-Altaic linguistic unity, resulting in the eastern extremity or extension of Figure 1 above, the small sector where **p-* and **x-* were, in one form or another, preserved; here also belong the earliest representatives of what later became pre- and Old Japanese.

4. The special linguistic-historical position of Manchu within the Tungus (and also the Altaic) languages is not because this language is 'typically' or 'representatively' southern, but rather because its earliest forms were originally involved spatially with the proto-Mongolian sector of Altaic, hence Manchu is 'typically southern' only in the special sense of once having been involved in the lower portion of that special zone where original **p-* was preserved in a variety of different ways, mostly in the form of dento-labials.

Common to all the above, and perhaps most important of the points that it is necessary to make in summation, is our concern for demonstrating — as we believe the present contribution has done — the methodological necessity for treating all elements in the history of the Tungus languages, including this immediate question of their original geographic distribution, in terms of the hypothesis of a still earlier Altaic linguistic unity, of which the Tungus family represents an early but secondary subsector. Any other hypothesis, and in particular the

hypothesis of a totally isolated Tungus family initially separate from Altaic, precludes even the possibility of significant linguistic-historical explanation for most of the linguistic phenomena now documented for this part of Asia, or for virtually all the developments that have transpired in the course of the history of these numerous non-Slavic languages of the USSR.

University of Washington

Notes

¹This, it is interesting to note, is the same I.-E. root whose borrowing into Ossetic from proto-Slavic was discussed at the Conference by Z. Gołab in his paper "Prehistoric contacts between Ossetic and Slavic."

²These brief translations are merely convenient semantic tags for the identification of the etyma concerned, and ought not to be mistaken for or misunderstood as in any way representing the 'basic meanings' of the forms in question.

References

- Benzing, Johannes. 1955. "Die tungusischen Sprachen, Versuch einer vergleichenden Grammatik." *Abhandlungen der geistes- und sozialwissenschaftlichen Klasse, Akademie der Wissenschaften und der Literatur, Mainz* 11:949-1099.
- Bloomfield, Leonard. 1933. *Language*. New York.
- Boodberg, Peter A. 1939. "Marginalia to the Histories of the Northern Dynasties." *HJAS* 4.230-83. Cited from the reprint in A. P. Cohen, ed. 1979. *Selected Works of Peter A. Boodberg*, Berkeley. pp. 296-349.
- Černjakov, Z. E. 1935. *Karta rasprostranjenija jazykov narodov Severa SSSR*. Moscow, Leningrad.
- Cincius, V. I. 1949. *Sravnitel'naja fonetika tunguso-man'čžurskix jazykov*. Leningrad.
- _____. et al. 1975, 1977. *Sravnitel'nyj slovar' tunguso-man'čžurskix jazykov*. Leningrad.
- Doerger, Gerhard. 1974. "Ist das Japanische mit den altaischen Sprachen verwandt?" *ZDMG* 124:1.103-42.
- _____. 1978. "Classification Problems of Tungus." *Tungusica* 1:1-26.
- Grube, Wm. 1900. *Linguistische Ergebnisse, 2, Goldische-Deutsches Wörterverzeichnis*, Appendix to Vol. 3, no. 2, in L. v. Schrenck. 1858. *Reisen und Forschungen im Amur-Lande in den Jahren 1854-1856...* St. Petersburg.
- Ikegami, J. 1974. "Versuch einer Klassifikation der tungusischen Sprachen." Pp. 271-2 in G. Hazai, ed., *Sprache, Geschichte und Kultur der altaischen Völker*. Berlin.
- _____. 1971. "Tsungūsū-go no henssen." Pp. 279-302 in Hattori, Shirō, ed., *Gengo no keitō to rekishi*. Tokyo.

- Ivanovskij, A. O. 1894. *Mandjurica, Obrazcy solonskogo i daxurskogo jazykov*. St. Petersburg.
- Janhunen, J. 1983. "Kita-Ajia no minzoku to gengo no bunrui." *Gekkan Gengo* 112:11 (= No. 142). 46-53.
- Menges, Karl H. 1968a. "Die tungusischen Sprachen." B. Spuler, ed., *Handbuch der Orientalistik, Erste Abteilung, Fünfter Bd., Dritter Abschnitt, Tungusologie*. Leiden: 21-256.
- _____. 1968b. *The Turkic languages and peoples, An Introduction to Turkic studies*. Wiesbaden.
- _____. 1975. *Altajische Studien, II. Japanisch und Altajisch*.
- Miller, Roy Andrew. 1971. *Japanese and the Other Altaic Languages*. Chicago.
- _____. 1980. *Origins of the Japanese Language*. Seattle.
- _____. 1981. "Altaic origins of the Japanese verb classes." Pp. 845-80 in Y. L. Arbeitman and A. R. Bomhard, eds., *Bono Homini Donum: Essays in Historical Linguistics, in Memory of J. Alexander Kerns*. Amsterdam: 845-80.
- _____. 1984. "Korean and Altaic." *Journal of Korean Studies*. 5:143-71.
- _____. 1985. "IE **ser-p-* 'sickle; hook' in Altaic." Pp. 93-122 in G. Jarring and S. Rosén, *Altaistic studies, Papers presented at the 25th Meeting of the PLAC at Uppsala June 7-11 1982*, Konferenser 12, Kungl. Vitterhets Historie och Antikvitets Akademien. Stockholm.
- _____. 1986. "Altaic Evidence for Prehistoric Invasions of Japan." *Ural-Altaische Jahrbücher* 58:39-61.
- _____. 1987. "Proto-Altaic **x-*." *Central Asiatic Journal* 31:1-2, 19-63.
- Poppe, N. N. 1931. *Materialy po solonskomu jazyku*. Leningrad.
- _____. 1960. *Vergleichende Grammatik der altaischen Sprachen, Teil 1, Vergleichende Lautlehre*. Wiesbaden.
- Schmidt, J. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar.
- TMS, see Cincius, V. I. et al., eds., 1975, 1977.

Observations on the Russian Component in Karelian

Stefan M. Pugh

This paper is a preliminary study of the Russian component in two dialects of Karelian spoken within the Soviet Union; the purpose of our analysis is to attempt to shed light upon the processes of change which take place in the language of a people that constitutes a small minority in a predominantly Russian-speaking region.¹

Before proceeding to the results of this study, it is worthwhile to discuss the status of Karelian: along with Finnish, Lude, Veps, Vote, Estonian, and Livonian, it is part of the Balto-Fennic branch of the Fenno-Ugric family of languages. There is no general agreement, however, as to whether Karelian should be considered a 'language' or a 'dialect'; B. Collinder, for example, identifies it as an 'Eastern Finnish dialect' (Collinder 1965:11) while Soviet studies often refer to Karelian as a 'language' (e.g., OFUJ 1975:6, Makarov 1969, Punžina 1971). We will not address that problem here. A major dialect variously treated by scholars is known as Olonets (or Olonetsian). In western scholarship, Olonets (Finnish *aunus*) is usually described as a variant or dialect of Karelian, and in fact one often finds the dialects of Karelian referred to collectively as 'Karelian-Olonets' (Collinder 1965:11, Oinas 1958:269). In Soviet works Olonets is also considered a dialect of Karelian (Russian *livvikovskij dialekt*); nevertheless it is considered to be different enough from other Karelian dialects to justify identifying the latter as 'Karelian proper' (*sobstvenno karel'skij jazyk*). In our study we will use the latter term in the same way, i.e., when differentiating between Olonets and other Karelian dialects, while the term 'Karelian' will refer to all dialects, including Olonets.

Karelian proper is primarily spoken in the northern regions of the Karelian ASSR (bordering on Finland), while Olonets is spoken in the south of the republic, near Petrozavodsk. In our paper the Tixvin dialect of Karelian (Tix.), which is spoken in the southeastern corner of the Leningrad *oblast'*, will be contrasted with Olonets (Ol.). The Tixvin dialect is described by Soviet scholars as a dialect of 'Karelian proper' (Rjagoev 1980) and was chosen for the analysis because its speakers have lived for so long (approximately three centuries) in the Russian language

environment, separated from the rest of Karelian. Olonets, on the other hand, was spoken for the most part on Finnish territory until the Winter War (1939-40), as were many Karelian dialects, and was purposely selected instead of a north Karelian dialect because it is described as 'strongly influenced' by Russian (Collinder 1965:11). The aim of this study is therefore to identify the manner and extent to which these two dialects have been influenced by Russian, and how (if at all) this influence differs in the two instances.²

I. Phonology. Our phonological analysis identified the presence of certain Russian features in both dialects. Thus, devoicing of voiced consonants finally and before voiceless consonants is found in some Russian lexical items: Ol. *sus'etka* (< Russ. *sosedka*) 'female neighbor', Tix. *vet'* (< Russ. *ved'*) 'after all, surely'. In both dialects palatalization of consonants before front vowels takes place in Russian as well as in native lexicon, and may be considered the most noteworthy aspect of Russian phonological interference. The presence of palatalized dental consonants before *i* or *j* in eastern Finnish, Karelian, Veps, Estonian, and Livonian has been pointed out before.³ In the texts analyzed, however, palatalization is found in consonants preceding not only *i* but *e* (*n'emeskoj* 'German', *üht'enä* 'one'), *ä* (*viel'ä* 'yet', *t'ämä* 'this'), *ö* and *ü* (*t'ürt'ö* 'girl').

Consonant clusters tend to be simplified in both dialects:

Tix. *spominan* < *vspominat'* 'remember'
n'emeskoj < *nemeckoj* 'German'
srojitti < *stroit'* 'build'

Ol. *klad* < *sklad* 'shed, storage place'
fat'eria < *kvartira* 'apartment'⁴

In Olonets we do find an unsimplified initial cluster *vst-* in *vstupi-* 'enter'; compare the simplification of *vsp-* > *sp-* in Tixvin. The former is evidently a recent borrowing in Olonets, and an example of the tendency to keep initial consonant clusters in newer loans.⁵

II. Lexicon. With reference to the lexical component, our study is concerned not so much with an enumeration of the substantives, verbs, etc., borrowed into Karelian from Russian; of greater interest, rather, is the question of how these elements fit into the native grammatical system (see 'Morphology' below). Many of the loan words have already been enumerated in other studies.⁶

Purely in terms of statistics, the number of Russian lexical items used in the two sets of texts was approximately equal in all categories except

the verb: many more verbs of Russian origin were identified in the Tixvin texts than in Olonets.

III. Morphology.

1. *The Noun.* In Olonets almost without exception Russian nouns are adapted to the native grammatical system. Thus the desinences denoting case and number are native Karelian, e.g., *voinan* 'war' (genitive), *kolhozas* 'kolhoz' (inessive), *roditel'at* 'parents' (nominative plural), *dorogal* 'road' (adessive); only two instances of the use of Russian desinences are found (out of fifty-seven in this dialect): *ovošši* 'vegetables' (nominative plural), and in the date *p'ervovo aprēl'ä* 'on April first' (genitive). In contrast, of sixty Russian borrowings found in the Tixvin texts, ten have Russian desinences, e.g., *kilōmetro*v 'kilometers' (genitive plural), *mog'ili* 'graves' (accusative plural), *voinu* 'war' (accusative singular),⁷ *kopejok* 'kopeks' and *nacional'nost'e*i 'nationalities' (both genitive plurals). The three genitive plurals cited occur after numerals, the use of which deserves mention here.

As in Finnish, the native Karelian system requires the partitive singular case after all numerals above 'one.' In the Olonets texts, this is precisely what we find. All the numerals used are of native origin, except the date cited above (*p'ervovo aprēl'ä*) and in the expression of years, in which Russian numerals are found: *sorok p'atoil vuvvel* 'in the year '45', and *d'evätnatcatoil vuvvel* 'in the year '19'. Nevertheless, the years are stated using native Karelian desinences.

We find more evidence of interference in the Tixvin system, however. When counting from one to ten, as well as by tens, Karelian numerals are used, with the partitive case (as expected): *kolme talua* 'three houses', *kūžikümmendä talua* 'sixty houses'. In the teens, both native and Russian elements are found: the native numeral plus the partitive case occurs four times (*kakšitoista talua* 'twelve houses'); the Russian numeral plus the Russian genitive plural occurs twice (*dvinatsat' kilometrov* 'twelve kilometers' and *pjatnatsat' kopejok* 'fifteen kopeks'); once we even find the native number plus the Russian genitive plural (*n'ellä nacional'nost'e*i 'four nationalities'). The last example is also of interest because one would expect the genitive singular with the number 'four,' according to the rules of Russian. The desinence *-a* found elsewhere (e.g., in *talua*) is to be identified as a native partitive rather than the Russian genitive singular.

It appears, therefore, that Tixvin has experienced considerable interference in the noun and numeral systems.

2. *The Verb.* It was mentioned above that there was a great difference in the number of borrowed verbs in the two sets of texts; in fact, more than twice the number as occurred in Olonets were found in Tixvin (14 and 38, resp.).

In the Olonets texts, every Russian verb occurs with a native Karelian desinence, e.g., *polučin* 'I received', *načnih* 'it began'. In contrast, it is noteworthy that there are three cases in Tixvin in which the entire Russian form is used: *kažotsa* 'it seems', *značit* '(that) means', and *roskazival* 'related'. The last occurs alongside the form *roskazivat*; since the desinence *-ivat* (3rd plural preterite in Finnish) does not exist in Karelian, this form has to be explained as another instance of interference, possibly on the model of the Russian present tense form *rasskazyvajut* (3rd plural).

Of special interest in the study of the Karelian verb is the use of verbs of motion. In both dialects studied here we find native and Russian verbs of motion: in Tixvin the ratio of native:Russian forms is 1:1, while in Olonets it turns out to be 20:1. In the latter dialect we find only two attestations of Russian verbs of motion (*perevedittih* 'they lead across', *pereidimö* 'we went across'), evidently used because of the special lexical meaning imparted by the prefix *pere-*; compare the use of unprefixated native *vein* 'I led'. In comparison, Tixvin makes use of four Russian prefixes: *pro-*, *pere-*, *pod-*, and *u-*. Tixvin thus seems to have acquired to some extent the Russian system of prefixation (absent in Balto-Fennic) in this verbal category, although the prefix may not always express the same meaning that it does in Russian. Thus, the prefix *u-* does not always express the concept of 'leaving,' rather it often corresponds to the use of *po-* in Russian (as a perfectivizer, marking the inception of an action). The use in Olonets of a prefixed Russian verbal form to express only a particular meaning is seen not only in the verbs of motion; compare the native Ol. verb *el'ie* 'to live', next to Russian *pereživiimö* 'we survived'.

As one might expect, given the highly developed system of verbal prefixation in Russian, the majority of verbs borrowed from Russian into both Olonets and Tixvin are prefixed. As our study shows, however, Tixvin appears to have assimilated the Russian system of prefixation to a greater extent than has Olonets.

3. *Conjunctions and Particles.* A comparison of the occurrence of Russian conjunctions and particles in the dialect texts also indicates patterns of greater assimilation by Tixvin than Olonets. In Tixvin the conjunctions *da*, *i*, *a* are very common; we find *ili* 'or' and *potomu što* 'because' once each; *što* 'that' and *bitto* 'as if' occur twice each. The particles *že* and *vet'* are found frequently; the definite particle *-to* (e.g.,

miššä-to 'somewhere') and the negative particle *n'i-* (*n'i-midä* 'nothing') both occur three times. In Olonets, we find *da*, *i* (but not *a*) many times, as in Tixvin; *l'ibo* 'or' and *jesl'i* 'if' occur once each. Of the particles, only *ni-* is found, occurring four times (e.g., *n'i-midä* 'nothing', *n'i-kedä* 'no one').

In conclusion, the data presented here suggest that long and uninterrupted contact of Tixvin with Russian has (as one might expect) brought about many changes in this dialect. Some of these are structurally significant, such as the use of Russian nominal and verbal desinences (alongside native desinences) and Russian verbal prefixes in borrowed lexicon. In spite of a relatively large number of Russian loanwords, in comparison the native grammatical structure of Olonets does not seem to have been strongly affected. The results of this study are of course only preliminary; many more texts will have to be analyzed before the changes indicated by our data can be confirmed as general tendencies. If they are so confirmed, it remains to be seen if future patterns of change in Olonets will resemble those we have seen in Tixvin.

Duke University

Notes

¹According to OFUJ 1975:6, in 1970 Karelians in the Soviet Union numbered 146,000.

²Since this is only a preliminary study, the texts used for our control data were relatively short. For both dialects the texts (totalling approximately 2500 words) were chosen according to similarity of content (e.g., history of a town, description of local customs) and age of the informants. Tixvin texts analyzed are: nos. 1 and 10 in Rjagoev 1977; no. 1 in Rjagoev 1980. Olonets texts: nos. 7, 11, 75, and 76 in Makarov 1969.

³OFUJ 1975:40. The role of Russian as a possible factor contributing to this development in the Balto-Fennic languages is not mentioned here.

⁴Cited as Ol. *fat'iera* (Russ. dialect *fat'era*) in Kalima 1955:55.

⁵J. Kalima points out that initial clusters of three consonants are rare, citing Kar.-Ol. *sproavie* 'improve' as an example where this is found (Kalima 1955:48).

⁶For example Kalima 1955 and Punžina 1971.

⁷If this is from Karelian *voina* (nom.) and not Ol. *voinu*, cited by Kalima (1955:143).

References

- Collinder, Björn. 1960. *Comparative Grammar of the Uralic Languages*. Stockholm: Almqvist and Wiksell.
- _____. 1965. *Introduction to the Uralic Languages*. Berkeley and Los Angeles: University of California Press.
- Kalima, Jalo. 1955. *Die slavischen Lehnwörter im Ostseefinnischen*. Berlin: Osteuropa-Institut an der freien Universität Berlin. Slavistische Veröffentlichungen. Band 8.
- Makarov, Grigorij Nikolaevič and Vladimir Dmitrievič Rjagoev. 1969. *Obrazcy karel'skoj reči. Govory livvikovskogo dialekta karel'skogo jazyka*. Leningrad: Nauka.
- OFUJ, 1975. = *Osnovy finno-ugorskogo jazykoznanija, pribaltijsko-finskie, saamskie i mordovskie jazyki*. Moscow: Nauka.
- Oinas, Felix J. 1958. "Russian and Eastern Balto-Finnic Linguistic Contact," *American Contributions to the Fourth International Congress of Slavists* 267-78. The Hague: Mouton.
- Punžina, A. V. 1971. "O nekotoryx leksičeskix zaimstvovanijax v karel'skom jazyke," *Pribaltijsko-finskoe jazykoznanie* 5:42-45.
- Rjagoev, Vladimir Dmitrievič. 1977. *Tixvinskij govor karel'skogo jazyka*. Leningrad: Nauka.
- _____. 1980. *Obrazcy karel'skoj reči (tixvinskij govor sobstvenno karel'skogo jazyka)*. Leningrad: Nauka.

The Scope of Saussure's Law in Colloquial Lithuanian

Steven Young

A non-native speaker of Lithuanian, like myself, who has mastered a linguistic description of the prosody of the standard language before gaining conversational fluency, cannot help but be struck by the frequent accentual departures from the norm in the colloquial language. Although I had been superficially aware of nonnormative accentuation in the speech of Chicago Lithuanians, who often acknowledged that they did not stress according to the rules they were taught, it was not until a prolonged research stay in Lithuania that I was able to observe the range of systematic innovations in accentual paradigms in the non-dialectal colloquial language.

Typologically, Lithuanian stress patterns are perhaps most reminiscent of Russian stress patterns, with a fixed stress paradigm and a shifting paradigm in which stress alternates between the stem and the desinence. Unlike Russian, most of the Lithuanian stress alternations result from a historical interplay of ictus and tone, which is still rather transparent in Lithuanian morphophonemics. This alternation is referred to in the literature as Saussure's Law. Saussure first showed that, for an earlier stage of Lithuanian in which so-called acute and circumflex tones were still preserved on now-unstressed long syllables, the primary ictus was attracted onto an acute syllable from an immediately preceding short or circumflex syllable, acute being in some sense the most prominent syllable in the word. Long acute-toned desinential syllables were later shortened and phonemic tone was lost outside the syllable with primary stress. In this connection, Saussure's Law lost its purely phonetic conditioning and began to function as a morphophonemic rule.

If the effects of Saussure's Law are reversed for the noun, though not for the verb, there will be a residue of stem/desinence alternations. This is the so-called mobile stress paradigm. For these alternations, inherited from Balto-Slavic, the prosodic composition of the stem plays no role — noun stems are simply marked in the lexicon as inherently fixed or mobile. But Saussure's Law can be superimposed on these stems, depending on the prosodic structure of the stem (short or circumflex syllable). This creates the traditional four accent classes of the

Lithuanian noun: 1 refers to fixed stem stress, 2 refers to fixed stress overlaid with Saussure's Law, 3 refers to mobile stress, and 4 to mobile stress overlaid with Saussure's Law.

Most of the deviations from the accentual norm in colloquial Lithuanian can be interpreted as either a suspension of Saussure's Law (that is, a generalization of stem stress) or an extension of Saussure's Law (that is, a generalization of stem/desinence stress) to new environments. There are a number of factors, phonetic and otherwise, which contribute to restricting the free application of Saussure's Law. The principle phonetic basis for this restricted application is the neutralization of the acute and circumflex tonal opposition on a long vowel (including the diphthongs *ie* and *uo*, which are functionally long vowels; the tonal opposition is still preserved on diphthongs). With the loss of the acute/circumflex distinction on long vowels (and thus the loss of the conditioning factor for Saussure's Law), stress patterns which had earlier depended on this opposition are no longer maintained. In the noun, given a stem in a long vowel, the 1st and 2nd stress patterns will merge as the 1st (*eĩnam į šokių* 'let's go to the dances') and the 3rd and 4th will also merge; for some speakers the third pattern is generalized at the expense of the fourth, while for others the fourth is generalized. The stress patterns are still preserved, however, for stems in a short vowel or a diphthong.

The direction of stress pattern generalization for verbal bases in a long vowel typically depends on the predominant etymological intonation of a particular formally or semantically defined subgroup. For many speakers, for example, present-tense stems ending in a soft consonant tend to generalize the circumflex pattern: in the first person singular, an environment for Saussure's Law, acute stem *košiù* 'filter, strain' behaves like *lošiù* 'play', a circumflex base; past *košiaũ*, *lošiaũ*. Present-tense stems ending in a hard consonant, on the other hand, tend to generalize the acute pattern, thus *miegu* 'I sleep' and *lieku* 'I remain', both circumflex bases. Nowhere in the nominal or verbal morphology are the two stress types maintained side by side in such stems for any formal subclass.

Failure of the tone-sensitive Saussure's Law to operate properly because of a loss of the acute/circumflex opposition does not, however, account for all the accentual discrepancies between the colloquial language and the norm. Some restrictions on the free application of Saussure's Law are nonphonologically motivated: stem stress is generalized in environments where Saussure's Law should apply, regardless of the length of the stem or intonation of the syllables involved.

Two subgroups can be distinguished. The first consists of nouns with polysyllabic bases. In place of end-stressed nom. sg. fem. *lašišà* 'salmon', *telegramà* 'telegram', *Aukštaitijà*, *Reginà*, in environments where Saussure's Law is expected, colloquial Lithuanian typically has *lašiša*, *telegrama*, *Aukštaitija*, *Regina*. Using the matrix sentence *aš pamačiaũ... 'I saw...'* for eliciting the accusative plural, informants gave me: *Aš pamačiaũ degtũkus* '...matches', *saldaĩnius* 'candies', *stiklines* 'glasses', *cigarėtes* 'cigarettes'; for instrumental singular, *su leidĩmu* 'with permission', *su degtine* 'with liquor', *su Birũte*; for locative singular masculine, *aš buvaũ universitėte* 'I was at the university', *institũte* '...institute', *antikvairiũte* '...used-book store'. The question arises: can this tendency towards stem stress be expressed in purely nonsegmental phonological terms, based on a critical number of syllables? The answer is no: polysyllabic verb stems show no evidence for this generalization of stem stress — *gyvenũ, gyvenĩ, gyvėna* 'live (1 sg., 2 sg., 3)'. Nor is there any general tendency towards stress retraction for short final syllables which are not desinences: thus the third person future *susipràs* 'will understand' (compare the noun *intrĩgas* 'intrigues (acc. pl.)'), and also uninflected forms: *pamažũ* 'slowly', *niekadà* 'never', *referĩ* 'referee'. The generalization of stem stress in a Saussure-Law environment must thus be limited to nouns, and the penult must be separated from the final syllable by a morpheme boundary. (Actually, even this is too broad a statement, since mobile-stress forms often show final stress: *parodà* 'exhibit', *atidũs* 'attentive', *Lietuvà* 'Lithuania'.) We thus have a limited tendency toward full stem stress in polysyllabic stems — only in nouns, and then only in the case of nonmobile stems. It is not clear to me what the original motivation for this generalized stem stress is, or why it does not extend beyond nouns (I will return to this question below); the earliest accented Lithuanian texts give no evidence for such a retraction.

The other category in which stem stress is generalized comprises etymologically nonnative noun stems, whatever their syllabic or prosodic structure. (Borrowed verb stems typically acquire a Lithuanian suffix and are thus out of the picture.) One minimal pair formed in this way is the native *sagà* 'button', where Saussure's Law operates, and the nonnative *sāga* 'saga', where the rule does not apply (dictionaries will give the same headword, *sagà*, for both). Near minimal pairs include native *kũras* 'fuel', loc. sg. *kurė*, but nonnative *biũras* 'office', loc. sg. *biũre*; and native *stiklas* 'glass', instr. sg. (*su*) *stiklũ*, but nonnative *ciklas* 'cycle', instr. sg. (*su*) *ciklu*. Clearly, the difference in accentual behavior is not connected with any formal property of the stem; but the only nonphonological feature distinguishing these stems is that *biũras* is a borrowing, while *kũras* is part of the native lexicon. It might be countered that certain etymologically

nonnative stems nevertheless tend to preserve end stress in the appropriate form, for example, *kavà* 'coffee', but these typically belong to the fourth accent class, and are therefore mobile. In general, very few borrowings have entered the mobile category; those which succeed gain, as it were, full Lithuanian citizenship — they are fully assimilated accentually, and should probably be considered as synchronically native, like the above *stiklas*, originally a borrowing corresponding to Common Slavic **stǫklo*.

This tendency toward fixed stem stress for nonnative stems must be fairly recent; as for stem-stressed polysyllabic bases, I could find no trace of it in the earliest accented texts. An apparent paradox arises: at some point a set of stems (earlier borrowings) ceased to undergo Saussure's Law, but only if they were somehow marked as nonnative. The shift is undoubtedly connected with the stress of the word in the source language. The great majority of *earlier* nonnative stems in Lithuanian are borrowings from Byelorussian, the longtime administrative language of the Grand Duchy of Lithuania, or from Polish, through the cultural and religious influence of Poland. In nominal borrowings, the stress of the nominative singular (usually coinciding with the oblique-stem stress) is typically retained in forms where Saussure's Law does not apply. Thus *kalėdos* 'Christmas', *gaspadōrius* 'proprietor', from Byelorussian *kaljady*, *haspadār*. Borrowings from Polish show some hesitation in stress placement when the nominative singular masculine of the original has a zero ending. Some of these borrowings reflect the stress of the Polish nominative singular, while others reflect the stress of the Polish oblique stem: *pópiežius* (*papież*) 'Pope', *rōtušė* (*ratusz*) 'town hall', but *kalendōrius* (*kalendarz*), *testameñtas* (*testament*).

Where Saussure's Law would apply (nominative singular feminine, for example), there is no surface discrepancy as long as the source word is end-stressed. This end stress is found only for a small number of monosyllabic stems borrowed from Byelorussian. It is, of course, completely lacking in Polish borrowings. Where the lending language stresses the penult (stem-final syllable), there arises a discrepancy between the stress of the word in the source language and in Lithuanian: *bonkà* (Pol *bańka*) 'jar', *alyvà* (BR *alíva*, Pol *oliwa*) 'oil'. The point of departure for the suppression of Saussure's Law here is the realignment of the surface stress in the Lithuanian forms to reflect more closely the stress of the source. Once the pattern was established, earlier borrowings felt to be nonnative by bilingual speakers (most likely Lithuanian/Polish speakers of the Wileńszczyzna), were apparently adjusted to fit this set. All later borrowings — even those with end stress in the lending language

— enter colloquial Lithuanian with stem stress: colloquial *grāfa* (*grafà* in the dictionaries) ‘column in a table’, from Russian *grafá*.

We have examined a range of stems in which Saussure’s Law, for one historical reason or another, no longer applies. Let us now briefly consider the forms where Saussure’s Law continues to operate, producing short stressed final syllables: +native monosyllabic stems and polysyllabic verb stems which do not contain \tilde{V} . With regard to susceptibility to Saussure’s Law these form a closed set — lengthy, perhaps, but finite. No new words conceivably could be added to the +native monosyllabic stems, nor will innovations among verbs ever be subject to the rule such that a short final syllable would result. A productive verbal suffix will either contain an acute diphthong (e.g., *-áuti* in *prezidentáuti* : *prezidentáuja* ‘be president’) or an etymologically acute \tilde{V} , with the penultimate stress pattern preserved: *kolektivizúoti* : *kolektivizúoju* ‘collectivize’, *protestúoti* : *protestúoju* ‘protest’, *intensyvėti* : *intensyvėja* ‘intensify’. Productive formations in *-inti*, which alone would be decisive, retain the stress of the base word: *kolektývinti* (*kolektývas* ‘collective’) : *kolektývinu* ‘collectivize’, *įmagnėtinti* (*magnėtas* ‘magnet’) : *įmagnėtinu* ‘magnetize’. The possibility of a neologism obtaining a stressed short desinence thus does not arise for verbs. If it did, we might speculate that here too Saussure’s Law would fail to apply.

It is precisely in those categories which do not form closed sets that we find Saussure’s Law failing to apply: in nominal derivation (resulting in polysyllabic stems) and in borrowings.

There is, in fact, one more productive category where the rule, as expected, fails to operate, but for which I have incomplete material: adaptations of common nouns as personal names, particularly feminine (the starting point here is apparently the stress of the vocative). Thus *rasà* ‘dew’ has provided the name *Rāsa*, instr. (*su*) *Rāsa*; similarly *su linù* ‘with linen’ and *su Linu* ‘with Linas (adaptation of *linas*)’ form an accentual minimal pair.

In effect, no new word entering Lithuanian, either as a borrowing or a native innovation, will ever be subject to Saussure’s Law. In this, the inflectional accentology of Lithuanian has drawn closer to its derivational accentology. In derivation, Saussure’s Law was also once regulated by the underlying prosodic structure of component morphemes. This morphophonemic alternation was eventually suppressed, and the stress of certain morphemes became generalized along formal or semantic lines. This suppression of the stress alternation associated with Saussure’s Law as now taken hold in inflection. The particular causes for this suppression are varied: loss of intonational oppositions on a long vowel, the possible influence of Polish penultimate stress, and the generalization

of vocative stress in proper names. The end result, however, is the same: for a large portion of the Lithuanian lexicon Saussure's Law has been replaced in the colloquial language by a generalization of stress on a particular syllable throughout the paradigm. Saussure's Law remains, but as a minor rule of the language, operating on a finite set of forms.

University of Maryland, Baltimore County

Other Books From Slavica

Ronelle Alexander: *The Structure of Vasko Popa's Poetry*.

American Contributions to the Eleventh International Congress of Slavists (Bratislava, 1993), Literature, Linguistics, Poetics, ed. Robert A. Maguire and Alan Timberlake.

American Contributions to the Tenth International Congress of Slavists, Sofia, September, 1988, Linguistics, ed. Alexander M. Schenker.

American Contributions to the Tenth International Congress of Slavists, Sofia, September, 1988, Literature, ed. Jane Gary Harris.

American Contributions to the Ninth International Congress of Slavists (Kiev 1983) Vol. 1: Linguistics, ed. Michael S. Flier.

American Contributions to the Ninth International Congress of Slavists, (Kiev 1983) Vol. 2: Literature, Poetics, History, ed. P. Debrecezeny.

American Contributions to the Eighth International Congress of Slavists, Vol 1: Linguistics and Poetics, ed. Henrik Birnbaum.

American Contributions to the Eighth International Congress of Slavists Vol. 2: Literature, ed. Victor Terras.

Patricia M. Arant: *Russian for Reading*.

Howard I. Aronson: *Georgian: A Reading Grammar*.

James Bailey: *Three Russian Lyric Folk Song Meters*.

Natalya Baranskaya: *Неделя как неделя Just Another Week*, ed. Lora Paperno, Natalie Roklina, and Richard Leed.

Adele Marie Barker: *The Mother Syndrome in the Russian Folk Imagination*.

John D. Basil: *The Mensheviks in the Revolution of 1917*.

Christina Y. Bethin: *Polish Syllables The Role of Prosody in Phonology and Morphology*.

Henrik Birnbaum & Thomas Eekman, eds.: *Fiction and Drama in Eastern and Southeastern Europe: Evolution and Experiment in the Postwar Period*.

Henrik Birnbaum and Peter T. Merrill: *Recent Advances in the Reconstruction of Common Slavic (1971-1982)*.

Other Books From Slavica

Marianna D. Birnbaum: *Humanists in a Shattered World: Croatian and Hungarian Latinity in the Sixteenth Century*.

F. J. Bister and Herbert Kuhner, eds.: *Carinthian Slovenian Poetry*.

K. L. Black, ed.: *A Biobibliographical Handbook of Bulgarian Authors*.

Ralph Bogert: *The Writer as Naysayer Miroslav Krleža and the Aesthetic of Interwar Central Europe*.

Marianna Bogojavlensky: *Russian Review Grammar*.

Rodica C. Botoman, Donald E. Corbin, E. Garrison Walters: *Îmi Place Limba Română/A Romanian Reader*.

Gary L. Browning: *Workbook to Russian Root List*.

Ranko Bugarski and Celia Hawkesworth, eds.: *Language Planning in Yugoslavia*.

Diana L. Burgin: *Richard Burgin A Life in Verse*.

R. L. Busch: *Humor in the Major Novels of Dostoevsky*.

Terence R. Carlton: *Introduction to the Phonological History of the Slavic Languages*.

Jozef Ciger-Hronský: *Jozef Mak* (a novel), translated from Slovak.

Julian W. Connolly & Sonia I. Ketchian, eds.: *Studies in Honor of Vsevolod Setchkarev*.

Andrew R. Corin: *The New York Missal: A Paleographic and Phonetic Analysis*.

Anna Lisa Crone and Catherine V. Chvany, eds.: *New Studies in Russian Language and Literature*.

Carolina De Maegd-Soëp: *Chekhov and Women: Women in the Life and Work of Chekhov*.

William W. Derbyshire: *A Basic Reference Grammar of Slovene*.

Dorothy Disterheft: *The Syntactic Development of the Infinitive in Indo-European*.

Per Durst-Andersen: *Mental Grammar Russian Aspect and Related Issues*.

Other Books From Slavica

Thomas Eekman and Dean S. Worth, eds.: *Russian Poetics*.

M. J. Elson: *Macedonian Verbal Morphology A Structural Analysis*.

M. S. Flier and R. D. Brecht, eds.: *Issues in Russian Morphosyntax*.

John M. Foley, ed.: *Oral Traditional Literature A Festschrift for Albert Bates Lord*.

John Miles Foley, ed.: *Comparative Research on Oral Traditions: A Memorial for Milman Parry*.

Richard Frucht, ed.: *Labyrinth of Nationalism, Complexities of Diplomacy Essays in Honor of Charles and Barbara Jelavich*.

Isidore Geld: *Dictionary of Omissions for Russian Translators with Examples from Scientific Texts*

Diana Greene: *Insidious Intent: An Interpretation of Fedor Sologub's The Petty Demon*.

Charles E. Gribble: *Reading Bulgarian Through Russian*.

Charles E. Gribble: *Russian Root List with a Sketch of Word Formation*.

Charles E. Gribble: *A Short Dictionary of 18th-Century Russian/ Словарик Русского Языка 18-20 Века*.

Charles E. Gribble, ed.: *Studies Presented to Professor Roman Jakobson by His Students*.

George J. Gutsche and Lauren G. Leighton, eds.: *New Perspectives on Nineteenth-Century Russian Prose*.

Morris Halle, ed.: *Roman Jakobson: What He Taught Us*.

Morris Halle, Krystyna Pomorska, Elena Semeka-Pankratov, and Boris Uspenskij, eds.: *Semiotics and the History of Culture In Honor of Jurij Lotman Studies in Russian*.

William S. Hamilton: *Introduction to Russian Phonology and Word Structure*.

Pierre R. Hart: *G. R. Derzhavin: A Poet's Progress*.

Michael Heim: *Contemporary Czech*.

Michael Heim, Z. Meyerstein, and Dean Worth: *Readings in Czech*.

Other Books From Slavica

Warren H. Held, Jr., William R. Schmalstieg, and Janet E. Gertz: *Beginning Hittite*.

Peter Hill: *The Dialect of Gorno Kalenik*.

M. Hubenova & others: *A Course in Modern Bulgarian*.

Martin E. Huld: *Basic Albanian Etymologies*.

Charles Isenberg: *Substantial Proofs of Being: Osip Mandelstam's Literary Prose*.

Roman Jakobson: *Brain and Language*

L. A. Johnson: *The Experience of Time in Crime and Punishment*.

S. J. Kirschbaum, ed.: *East European History (Selected Papers from the Third World Congress for Soviet and East European Studies)*.

Emily R. Klenin: *Animacy in Russian: A New Interpretation*.

Andrej Kodjak, Krystyna Pomorska, and Kiril Taranovsky, eds.: *Alexander Puškin Symposium II*.

Andrej Kodjak: *Pushkin's I. P. Belkin*.

Andrej Kodjak, Michael J. Connolly, Krystyna Pomorska, eds.: *Structural Analysis of Narrative Texts*.

Demetrius J. Koubourlis, ed.: *Topics in Slavic Phonology*.

Mark Kulikowski: *A Bibliography of Slavic Mythology*.

Konstantin Kustanovich: *The Artist and the Tyrant: Vassily Aksenov's Works in the Brezhnev Era*.

Ronald D. LeBlanc: *The Russianization of Gil Blas: A Study in Literary Appropriation*.

Richard L. Leed, Alexander D. Nakhimovsky, and Alice S. Nakhimovsky: *Beginning Russian, Second Revised Edition*.

Richard L. Leed and Slava Paperno: *5000 Russian Words With All Their Inflected Forms: A Russian-English Dictionary*.

Edgar H. Lehrman: *A Handbook to Eighty-Six of Chekhov's Stories in Russian*.

Lauren Leighton, ed.: *Studies in Honor of Xenia Gąsiorowska*.

Other Books From Slavica

Gail Lenhoff: *The Martyred Princes Boris and Gleb: A Social-Cultural Study of the Cult and the Texts.*

Jules F. Levin and Peter D. Haikalis, with Anatole A. Forostenko: *Reading Modern Russian.*

Maurice I. Levin: *Russian Declension and Conjugation: A Structural Description with Exercises.*

Alexander Lipson: *A Russian Course.*

Yvonne R. Lockwood: *Text and Context Folksong in a Bosnian Muslim Village.*

A Sense of Place Tsarskoe Selo and Its Poets Papers from the 1989 Dartmouth Conference Dedicated to the Centennial of Anna Akhmatova, ed. Lev Loseff and Barry Scherr.

Horace G. Lunt: *Fundamentals of Russian.*

Thomas G. Magner, ed.: *Slavic Linguistics and Language Teaching.*

Robert Mann: *Lances Sing: A Study of the Igor Tale.*

Stephen Marder: *A Supplementary Russian-English Dictionary.*

V. Markov and D. S. Worth, eds.: *From Los Angeles to Kiev Papers on the Occasion of the Ninth International Congress of Slavists.*

Cynthia L. Martin, Joanna Robin, and Donald K. Jarvis: *The Russian Desk: A Listening and Conversation Course.*

Mateja Matejić and Dragan Milivojević: *An Anthology of Medieval Serbian Literature in English.*

Peter J. Mayo: *The Morphology of Aspect in Seventeenth-Century Russian (Based on Texts of the Smutnoe Vremja).*

Arnold McMillin, ed.: *Aspects of Modern Russian and Czech Literature*

Gordon M. Messing: *A Glossary of Greek Romany As Spoken in Agia Varvara (Athens).*

Vasa D. Mihailovich and Mateja Matejic: *A Comprehensive Bibliography of Yugoslav Literature in English, 1593-1980.*

Other Books From Slavica

Dragan Milivojević and Vasa D. Mihailovich: *A Bibliography of Yugoslav Linguistics in English 1900-1980*.

Edward Mozejko, ed.: *Vasilij Pavlovich Aksënov: A Writer in Quest of Himself*.

Edward Mozejko: *Yordan Yovkov*.

Alexander D. Nakhimovsky and Richard L. Leed: *Advanced Russian, Second Edition, Revised*.

Wiesław Oleksy and Oscar E. Swan: *W Labiryncie (Labyrinth of Life) opera mydlana w dwudziestu jeden odcinkach oparta na motywach serialu telewizyjnego w reżyserii Pawła Karpińskiego* (a video-based advanced Polish language course).

Hongor Oulanoff: *The Prose Fiction of Veniamin Kaverin*.

T. Pachmuss: *Russian Literature in the Baltic between the World Wars*.

Lora Paperno: *Getting Around Town in Russian: Situational Dialogs*, English translation and photographs by Richard D. Sylvester.

Slava Paperno, Alexander D. Nakhimovsky, Alice S. Nakhimovsky, and Richard L. Leed: *Intermediate Russian: The Twelve Chairs*.

Ruth L. Pearce: *Russian For Expository Prose*.

Jan L. Perkowski: *The Darkling A Treatise on Slavic Vampirism*.

Gerald Pirog: *Aleksandr Blok's Итальянские Стужи Confrontation and Disillusionment*.

Leonard A. Polakiewicz: *Supplemental Materials for First Year Polish*.

Stanley J. Rabinowitz: *Sologub's Literary Children: Keys to a Symbolist's Prose*.

Gilbert C. Rappaport: *Grammatical Function and Syntactic Structure: The Adverbial Participle of Russian*.

Don Karl Rowney, ed.: *Imperial Power and Development: Papers on Pre-Revolutionary Russian History*

Don K. Rowney & G. Edward Orchard, eds.: *Russian and Slavic History*.

Other Books From Slavica

Klaus Roth and Gabriele Wolf, eds.: *South Slavic Folk Culture A Bibliography*.

Catherine Rudin: *Aspects of Bulgarian Syntax: Complementizers and WH Constructions*.

Norma L. Rudinsky: *Incipient Feminists: Women Writers in the Slovak National Revival*.

Gerald J. Sabo, S.J., ed.: *Valaská škola*, by Hugolin Gavlovič, with a linguistic sketch by Ľubomír Ďurovič.

Barry P. Scherr and Dean S. Worth, eds.: *Russian Verse Theory*.

William R. Schmalstieg: *Introduction to Old Church Slavic*.

R. D. Schupbach: *Lexical Specialization in Russian*.

P. Seyffert: *Soviet Literary Structuralism: Background Debate Issues*.

Kot K. Shangriladze and Erica W. Townsend, eds: *Papers for the V. Congress of Southeast European Studies (Belgrade, September 1984)*.

J. Thomas Shaw: *Pushkin A Concordance to the Poetry*.

J. Thomas Shaw: *Pushkin's Poetry of the Unexpected: The Nonrhymed Lines in the Rhymed Poetry, and The Rhymed Lines in the Nonrhymed Poetry*.

Efraim Sicher: *Style and Structure in the Prose of Isaak Babel'*.

Rimvydas Šilbajoris: *Tolstoy's Aesthetics and His Art*.

M. S. Simpson: *The Russian Gothic Novel and its British Antecedents*.

David A. Sloane: *Aleksandr Blok and the Dynamics of the Lyric Cycle*.

Greta N. Slobin, ed.: *Aleksej Remizov: Approaches to a Protean Writer*.

Theofanis G. Stavrou and Peter R. Weisensel: *Russian Travelers to the Christian East from the Twelfth to the Twentieth Century*.

G. Stone and D. S. Worth, eds.: *The Formation of the Slavonic Literary Languages, Proceedings of a Conference Held in Memory of Robert Auty and Anne Pennington at Oxford 6-11 July 1981*.

Other Books From Slavica

John W. Strong, ed.: *Essays on Revolutionary Culture and Stalinism*

Oscar E. Swan and Sylvia Gálová-Lorinc: *Beginning Slovak.*

Oscar E. Swan: *First Year Polish.*

Oscar E. Swan: *Intermediate Polish.*

Charles E. Townsend: *Continuing With Russian.*

Charles E. Townsend: *Czech Through Russian.*

Charles E. Townsend: *A Description of Spoken Prague Czech.*

Charles E. Townsend: *Russian Word Formation.*

Janet G. Tucker: *Innokentij Annenskij and the Acmeist Doctrine.*

Boryana Velcheva: *Proto-Slavic and Old Bulgarian Sound Changes.*

Walter N. Vickery, ed.: *Aleksandr Blok Centennial Conference.*

Essays in Honor of A. A. Zimin, ed. D. C. Waugh.

Daniel C. Waugh: *The Great Turkes Defiance On the History of the Apocryphal Correspondence of the Ottoman Sultan in its Muscovite and Russian Variants.*

Paul Wexler: *The Ashkenazic Jews: A Slavo-Turkic People in Search of a Jewish Identity.*

James B. Woodward: *Form and Meaning: Essays on Russian Literature.*

James B. Woodward: *The Symbolic Art of Gogol: Essays on His Short Fiction.*

Dean S. Worth: *Origins of Russian Grammar Notes on the state of Russian philology before the advent of printed grammars.*

Yordan Yovkov: *The Inn at Antimovo and Legends of Stara Planina*, translated from Bulgarian by John Burnip.

Что я видел What I Saw by Boris Zhitkov, annotated and edited by Richard L. Leed and Lora Paperno.

Twelve Stories by M. Zoshchenko, selected and annotated for English-speaking students by Lesli LaRocco and Slava Paperno.

Other Books From Slavica

JOURNALS

The International Journal of Slavic Linguistics and Poetics.

Folia Slavica

Oral Tradition

Balkanistica: Occasional Papers in Southeast European Studies